Neural computation in small sensory systems

Lessons on sparse and adaptive coding

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)
im Fach Biologie

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
Humboldt-Universität zu Berlin

von
Herrn Jan Clemens, Diplombiologe
geboren in Berlin

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Dr. Andreas Herrmann

Gutachter:
1. Prof. Dr. Bernhard Ronacher
2. Prof. Dr. Jan Benda
3. Prof. Dr. Martin Nawrot

eingereicht am: 27.03.2012
Tag der mündlichen Prüfung: 26.06.2012

Abstract

The goal of computational neuroscience is to describe the stimulus transformations performed by neural systems and to elucidate their mechanisms and functions. This thesis combines experiment, data analysis and theoretical modeling to explore neural coding in the small auditory systems of grasshoppers and crickets.

The first part deals with the transformation of the neural representation of courtship signals in grasshoppers. The code in auditory receptors is relatively homogeneous. That is, all neurons represent a very similar stimulus feature. Representation in higher-order neurons leads to an increase of temporal and population sparseness. This creates a labeled-line population code where different neurons represent different and specific stimulus features. Sparseness in the system increases through a nonlinear combination of two stimulus features. This transformation enables a simple mode of pattern classification, which ignores the timing of individual features and relies only on their average values during a signal. The transformation can therefore facilitate the recognition of the long, temporally redundant communication signals produced by grasshoppers and other insects.

The second part shows that spectral and temporal tuning of second-order neurons in crickets strongly depends on the complexity of the stimulus. While tuning is relatively broad for single-carrier stimuli, signals containing multiple carrier frequencies lead to a sharpening of the tuning. This sharpening preserves information about individual components of a complex stimulus. A network model revealed that such adaptive tuning can be implemented in a static network with mechanisms that are ubiquitous in many neural systems.

In summary, this study shows that the nervous systems of insects combine a relatively simple structure with complex stimulus transformations. This renders them empirically accessible and suitable model systems for computational neuroscience.

Zusammenfassung

Das Ziel von *computational neuroscience* ist, neuronale Transformationen zu beschreiben und deren Mechanismen und Funktionen zu beleuchten.Diese Doktorarbeit kombiniert Experiment, Datenanalyse und theoretische Modelle um neuronale Kodierung anhand des auditorischen Systems von Feldheuschrecke und Grille zu erforschen.

Der erste Teil befasst sich mit der Umwandlung neuronaler Repräsentation von Balzsignalen in Feldheuschrecken. In auditorischen Rezeptoren ist die Kodierung dieser Signale relativ homogen — alle Neuronen bilden den Reiz gleich ab. In nachgeschalteten Zellen wird die Kodierung spärlicher, sowohl auf Ebene der Zeit als auch der Zellpopulation. Es entsteht dabei ein so genannter *labeled line code*, bei dem unterschiedliche Nervenzellen unterschiedliche, spezifische Merkmale des Stimulus abbilden. Dieser Transformation liegt eine nichtlineare Kombination von mehreren Stimulusmerkmalen zu Grunde.

Die erhöhte Spezifizität von Neuronen dritter Ordnung ermöglicht eine einfache Art der Musterklassifikation, bei der die Zeitpunkte bestimmter Reizelemente innerhalb des Signals ignoriert werden können. Die beschriebene Reiztransformation repräsentiert einen möglichen Mechanismus für die Erkennung zeitlich redundanter Kommunikationssignale, wie sie von Feldheuschrecken und anderen Insekten produziert werden.

Im zweiten Teil wird gezeigt, dass die spektrale und zeitliche Abstimmung von Neuronen zweiter Ordnung bei Grillen von der Komplexität des Reizes abhängt. Während die Abstimmung für Reize mit nur einer Trägerfrequenz relativ breit ist, führen Reize mit mehreren Trägerfrequenzen zu einer Schärfung. Hierdurch kann Information über einzelne Komponenten eines komplexen Signals in der Kodierung erhalten werden. Ein statisches Netzwerkmodell zeigt, dass diese adaptive Abstimmung mit Mechanismen erzeugt werden kann, die in Nervensystemen vieler Organismen vorkommen.

Wie diese Doktorabeit zeigt, vereinen Insekten einfach aufgebaute und daher empirisch gut zugängliche Nervensysteme mit unerwartet komplexen Reiztransformationen. Dies macht sie zu nützlichen und produktiven Modellorganismen für die Neurowissenschaften.

Contents

1. Introduction

Nervous systems equip animals with the ability to react to the environment in an adaptive manner. We are used to describe the action of nervous system as "computation". Understanding neural computation entails three tasks:

1. To describe the *transformations* a nervous system performs on representations of the environment.

2. To elucidate the *mechanism* by which these transformation are implemented in neuronal hardware.

3. To determine what *functions* these transformations serve.

This thesis approaches these tasks in two examples, sparse coding and stimulus-dependent coding, using the auditory systems of insects as models. In the following, the model systems and the questions covered in this thesis will be introduced.

1.1. Using "toy" systems

This thesis will use the auditory system of insects to understand principles of neural computation because these relatively small and simple neuronal networks are highly accessible—not only experimentally but also conceptually. While animals with larger brains appear to exhibit more complex behavior (but see Chittka and Niven (2009)), they are not perfectly well suited to study the neural basis of this behavior. A nervous system's large size can lead to a great degree of redundancy and distribution of neural computations, making it hard to localize individual circuits controlling a given behavior. This is further complicated by their immense developmental plasticity which leads to each individual being equipped with idiosyncratic neural networks—at least on a fine spatial scale.

In contrast, the early stages of insect auditory systems consist of relatively few and identified cell types with genetically-fixed physiology and connectivity. This facilitates the integration of data across individuals. As the network structure is relatively simple, the amount of possible mechanisms underlying a given behavior is more restricted than in large and complex cortical networks. Furthermore, the task of insect auditory systems is fairly well-defined and restricted: The acoustic sense is used for the detection of predators and for intra-specific communication using genetically-fixed signals. The fact that the insect auditory systems studied in this thesis are used for specific and important tasks suggests a great degree of co-adaptation between the signals and the neural hardware that processes them. This tight connection between the nervous system and a specific behavior facilitates

functional explanations of neural computations. In the following, the structure and function of the two model systems used in this thesis will be described: the auditory system of grasshoppers and of crickets.

1.1.1. Grasshoppers

Work on the species *Chorthippus biguttulus* has provided valuable insights into the neural basis of song recognition in the early auditory system of acridid grasshoppers. On hot summer days, males produce a calling song by rubbing their hindlegs against a hardened vein on their forewings. This produces a broadband sound with power between 4 and 40 kHz. Species-specific motor programs yield a movement pattern that modify this carrier with a species- and sex-specific amplitude modulations (Elsner, 1974; von Helversen and von Helversen, 1997). In the species *Chorthippus biguttulus*, this envelope consists of 20–30 repetitions of a basic sub-unit—a loud pulse followed by a shorter and much softer pause (von Helversen, 1972). If a female grasshopper of the same species hears this song and considers it attractive, it responds with her own song. This female response allows the male to localize and approach the female, resulting eventually in copulation (von Helversen and von Helversen, 1987; Schul et al., 1999).

Early steps of auditory processing are well understood (Stumpner and Ronacher, 1991; Stumpner et al., 1991; Machens et al., 2001; Wohlgemuth and Ronacher, 2007; Gollisch et al., 2002; Vogel et al., 2005; Vogel and Ronacher, 2007; Neuhofer et al., 2008): The ears of the animals lie bi-laterally in the first abdominal segment. From here, 60–80 receptors per side project to a small three-layer feed-forward network in the animal's meta-thoracic ganglion. In addition to the receptors as an input layer, the network consists of an intermediate layer of 10–15 local interneurons and an output layer of another 15–20 ascending neurons. These in turn project to the brain where the final evaluation of song takes place. So far, only within-layer, but no recurrent connections, have been described, rendering this network structurally similar to a classic perceptron with a hidden layer (Boyan, 1999; Vogel et al., 2005; Vogel and Ronacher, 2007).

The connectivity of the network as well as the morphology and physiology of local and ascending neurons is preserved across individuals. Intracellular recording followed by staining of the cells allows one to identify morphologically-defined cell types (Römer and Marquart, 1984). The large diversity of neurons combined with the fact that they are identified, renders this system amenable to study population coding in subsequent layers. Interestingly, the structure of the network seems to be conserved not only within a species but also within the whole taxon, as comparative studies between the migratory locust *Locusta migratoria* and its rather distant relative *Chorthippus biguttulus* have found high similarity in dendritic and axonal morphology of identified cell types (Stumpner, 1988). Accordingly, the response properties of the network are conserved in both species (Neuhofer et al., 2008). This is especially interesting as locusts do not seem to rely as heavily on acoustic communication for sex than do species from the *Chorthippus* group (but see Dörscheidt and Rheinlaender (1980); Pflüger and Field (1999); Kalmring (1975)).

1.1.2. Crickets

In contrast to the relatively large diversity of cell types in grasshoppers, the early auditory system of the cricket can be reduced to a small, three-neuron network consisting of one local inhibitory interneuron Omega 1 (ON1) and two output or ascending neurons AN1 and AN2. Interestingly, the network implements a common network motif: feed-forward inhibition. There, however, exist other cell types whose computational role is so far little understood (Wohlers and Huber, 1982). All three "major" neurons can be relatively easily recorded extracellularly and identified either by the relative height of their action potentials or based on their response properties (Hennig, 1988; Marsat and Pollack, 2004). This allows the experimenter to get stable long-term recordings of the three neurons.

The task of this three-neuron network is two-fold: recognition and localization of mating signals, enabling positive phonotaxis, and detection and localization of predator signals, allowing evasive behavior (Nolen and Hoy, 1986a; Wyttenbach et al., 1996). Both signals are discriminated by the cricket on the basis of carrier frequency: mating signals usually exhibit a species-specific carrier between 3 and 6 kHz (Kostarakos et al., 2009, 2008); cricket-hunting bats can be recognized by their ultra-sound echolocation signals. This partition of the signal spave nicely maps to the structure of the auditory system: There exist two receptor populations, which are tuned to the carrier frequency of mating and echolocations signals, respectively (there exists a third population of mid-frequency receptors whose function is so far unknown, Imaizumi and Pollack (1999, 2001)). The two ascending neurons are driven most strongly by either receptor population. The ascending neuron AN1 receives input from low-frequency receptors and is involved in the processing of mating signals; male crickets produce calling songs with a species-specific carrier frequency and envelope, which is used by females to recognize, localize and approach males (Hennig, 1988, 2009). In contrast, the ascending neuron AN2 is most sensitive to high carrier frequencies associated with echolocations signals of cricket-hunting bats (Marsat and Pollack, 2006, 2010). AN2 has been shown to be necessary and sufficient to elicit avoidance behavior in response to ultra sound (Nolen and Hoy, 1984).

The inhibitory local neuron ON1 is sensitive at low and high carrier frequencies and inhibits both AN (Pollack, 1994; Selverston et al., 1985). The omega neuron inhibiting the AN on one side of the animal (ipsilateral) receives inputs form receptor populations of the other side's ear—the contralateral ear. This (contra-)lateral inhibition has been shown to enhance the lateral contrast and to improve localization of a sound source for positive and negative phonotaxis (Schildberger, 1988; Marsat and Pollack, 2005). However, as has been shown for other implementations of feed-forward inhibition, it also influences the temporal selectivity of both AN (Tunstall and Pollack, 2005).

The fact that the function of both AN can be mapped by their sensitivity to a simple, ecologically relevant signal parameter—the carrier frequency—together with the complication of both AN receiving the same, broadly tuned inhibitory signals allows one to investigate how different, concurrently presented signals are encoded by a small network.

1.2. The two big questions - sparse coding and stimulus-dependent coding

Having introduced both model systems, I will now expose the two topics this thesis deals with: sparse coding and stimulus-dependent coding. These questions will be motivated and their relevance for understanding neural computation in general and the auditory system of grasshoppers and crickets in particular will be outlined.

1.2.1. Sparse coding

A sparse code is a code which transmits as much information as possible with as few spikes as possible. However, sparse coding is not simply a code with few spikes, that is, with a low average firing rate. Indeed, common measures of sparseness are orthogonal to firing rate (Willmore and Tolhurst, 2001; Willmore et al., 2011). Rather, they signify a special, higher-order property of neural responses: their "clusteriness" (Conor Houghton, personal communication). Usually, two types of sparseness are discriminated. A *temporally sparse* response is one in which spikes come in packets and are interleaved by long stretches of silence. A *population sparse* response is one were few neurons in an assembly are active at the same time.

Sparse codes have been found in many neural systems, spanning animal taxa and sensory modalities, and have been suggested to have many advantages (see e.g. (Olshausen and Field, 2004; Vinje and Gallant, 2000; Papadopoulou et al., 2011; Chacron et al., 2011; Ohiorhenuan et al., 2010; Ito et al., 2008)). In a sparse code, firing is confined to specific events in time or in a population. Rare and synchronous population events or burst-like firing patterns are much more efficient in driving a downstream neuron (Wang et al., 2010). Prolonged weak responses would be much less effective in eliciting output spikes due to synaptic depression or spike-frequency adaptation (Benda et al., 2001). Furthermore, near-binary, sparse responses render differences between different inputs more salient as compared to a weak modulation of the instantaneous firing rate. This also increases the signal-to-noise ratio. However, it can also yield wasteful codes and lead to an inefficient exploitation of neuronal bandwidth, as the "codebook" employed by an ultrasparse neuron with binary responses is reduced to only two states, `ON` and `OFF` (Barlow, 2001).

In the retina, sparse coding is thought to serve compression of the neural representation: the decorrelation of feature selectivity reduces redundancy between ganglion cells and enables the transmission of as much information as possible through the information bottleneck of the optic nerve (Barlow, 2001; Zhaoping, 2006). In higher sensory areas like primary visual cortex V1 or the mushroom body in the olfactory systems of insects, compression is an unlikely objective, as these networks are highly divergent, exhibiting much more output than inputs axons. Here, representations are expanded and increase their specificity. Such sparse, overcomplete codes are thought to facilitate subsequent computations (Olshausen and Field, 2004; Zhaoping, 2002; Perez-Orive et al., 2002). However, few connections between a sparse code and its role in a specific behavioral context have been made. Sparseness is hypothesized to underlie rapid and efficient learning of odors in the olfactory system of

locusts (Stopfer et al., 1997; Ito et al., 2008; Leonardo, 2005). Also, sparseness has been implied in the computation of saliency of visual features in a scene (Zhaoping, 2002). Insect auditory systems with their restricted set of tasks might inspire hypotheses that add to this list. Indeed I will present data that suggests a role of sparse coding in enabling a simple algorithm for song recognition in the grasshopper (see chapter 4).

How sparse representations emerge in nervous systems is also not fully understood. Among the computations and biophysical mechanisms that have been reported to underlie sparseness are feedforward, feedback, lateral or recurrent inhibition, oscillations, spike-frequency adaptation or a high firing threshold (Perez-Orive et al., 2002; Papadopoulou et al., 2011; Rozell et al., 2008; Poo and Isaacson, 2009; Kapfer et al., 2007; Farkhooi et al., 2009). Note, that all these computations are not mutually exclusive and one can lead to the emergence of another (Stopfer et al., 1997). Revealing abstract computations underlying sparseness in the grasshopper might provide a more intuitive understanding of how sparseness is created (see chapter 4).

While sparse coding seems to be a widespread phenomenon, it has been reported mostly for large systems—that is, large relative to the auditory system of the grasshoppers. In part 1 of this thesis, I will examine sparse coding in the auditory system of the grasshopper, covering its description (chapter 2), generation (chapter 3) and putative function (chapter 4). The advantages of the auditory system of the grasshopper will be exploited: First, the small number of cells facilitates the relatively exhaustive characterization of responses of the whole network to natural stimuli with respect to temporal and population sparseness. Second, the relatively simple connectivity scheme reduces the number of possible mechanisms yielding sparse codes. And, third, the system's restricted task will allow me to propose a function for sparse coding in the grasshopper.

1.2.2. Stimulus-dependent coding

The second topic covered in this thesis is stimulus-dependent or adaptive coding. The phenomenon of adaptation is among the first ones students of neurophysiology are teached: if one moves from a dark to a bright room, at first one is blinded—everything appears to be white. After a while, a multitude of adaptive mechanisms kicks in to set the sensitivity of our visual system such that one is able to resolve differences in the visual scene. In humans, this is mostly done by correcting for the mean (luminance) and variance (contrast) of the current brightness distribution in the retina and the primary visual cortex, respectively (Carandini and Heeger, 2011). Information theory has shown, that this is actually optimal if one is interested in the spatial or temporal fluctuation of brightness (Laughlin, 1981; Sharpee et al., 2006; Fairhall et al., 2001)). However, adaptation also entails a trade-off: by normalizing for the mean and variance of the stimulus distribution the nervous system looses potentially important information about these aspects of the environment (Fairhall et al. (2001); Seriès et al. (2009); Hildebrandt (2010), but see Lundstrom and Fairhall (2006)).

While the mechanisms and functions of adaptation are fairly well understood in such cases, natural sensory environments vary in aspects other than the mean or the variance: In the auditory environment of temperate forests, a singing cricket faces much less competition

for signal space than in a tropical rainforest (Schmidt et al., 2011). Here, a female cricket localizing a singing male has to deal with little backgroud noise. However, in the tropics, signals produced by a multitude of other insects fill the spectrum up to the ultrasound range, posing a challenge for the auditory system which has to separate signal and background.

The fact that a given signal, as defined by its carrier frequency, can occur in relative isolation or in a background of other carrier frequencies poses a challenge to sensory systems, especially if tuning for carrier frequency is relatively broad and the background is as loud as the signal. In such cases, task-, context- or stimulus-dependent codes have been reported in mammalian auditory systems (Fritz et al., 2003; Ahrens et al., 2008; Schneider and Woolley, 2011). While the mechanism and purpose of adaptive coding is well understood for the simple cases of adaptation to the mean or the variance of a stimulus, the specific function of stimulus-dependent coding and its mechanisms are still elusive (but see Fritz et al. (2003)). Again, the simple structure and restricted task of insect auditory systems will help to shed light on these issues.

In part II of this thesis, stimulus-dependent coding in the cricket will be characterized (chapter 5). First, the relatively simple structure and semantics of the sensory environment of crickets—low carrier frequencies promise sex and high carrier frequency mean death—allows one to predict how adaptive coding should act and what possible benefits it might have. I will, second, test whether these hypotheses are true in the cricket using experimental data. Lastly, the simple structure of the early auditory system of the cricket, which, however, contains a universal network motif—feed-forward inhibition—will help reveal putative mechanisms underlying stimulus-dependent coding.

1.3. The structure of this thesis

Part I will be concerned with sparse coding in grasshoppers. In chapter 2, I will characterize temporal and population sparseness, describe the transformation of the population code for song in different stages of the network. I will derive hypotheses about possible mechanisms and benefits of increased sparseness and specificity at the network's output and elaborate on these in the following two chapters: Chapter 3 will show that two kinds of nonlinear computations underlie the emergence of sparseness in the grasshopper and discuss possible biophysical implementations. In chapter 4, I will train a classifier, inspired by the structure and function of the early auditory system of the grasshopper, on behavioral data to provide evidence that the emergence of a sparse and specific representation in the grasshopper enables a simple algorithm for song recognition.

Part II, which contains only a single chapter, will discuss the existence, benefits and mechanisms of stimulus-dependent, adaptive coding in the auditory system of the cricket. Here, I will characterize the dependence of encoding properties on stimulus complexity. A rate-based network model of the early auditory system of the cricket will suggest putative mechanisms underlying stimulus-dependent coding.

As the different chapters apply specifics methods for data collection and analysis, each chapter will be relatively self-contained with its own general introduction, methods, results and discussion. The results will be summarized in a final concluding chapter.

A lot of the results presented in this thesis would have been impossible without the data provided by others. Sandra Wohlgemuth provided most of the data analyzed in chapter 2 (recordings of local and ascending neurons). Olaf Kutzki contribute additional recordings of auditory receptors. The classifier in chapter 4 could not have been tested without the behavioral data kindly provided by Nicole Stange, Olaf Kutzki, Arne Schmidt and Jana Sträter. The data analyzed in chapter 5, were recorded by Florian Rau. Viktor Naumov performed recordings which are not included in this thesis but which served as pre-tests.

Part I.

Sparse coding in the auditory system of the grasshopper

2. Efficient transformation of an auditory population code in a small sensory system

This chapter will describe the emergence of sparseness in the early auditory system of the grasshopper. It will show the consequences of this transformation for the population code and discuss its benefits for the behavioral evaluation of communication signals.

2.1. Introduction

To increase their fitness, most animals strive to evaluate sensory signals that reveal the quality of a potential mate. What if an animal has only a few dozen neurons to preprocess this extremely important information? Optimal coding theory suggests that the creation of a sparse, decorrelated representation would be a wise investment of scarce neuronal resources (Barlow, 2001). That is indeed what has been found in many sensory modalities and species under natural conditions (Vinje and Gallant, 2000; DeWeese et al., 2003; Laurent, 2002).

These early sensory networks often comprise large numbers of cells and organize information in a map-like fashion, where spatial proximity of neurons reflects similarity in the selectivity for fundamental stimulus features (Yaksi et al., 2009; Kaschube et al., 2010). These maps tend to have a complete representation of sensory space and enable subsequent processing steps to select relevant features based on attention or associative learning. Reading out such a representation by "blind" summation of responses across different neurons would be highly inefficient to recover information because stimulus features are not only encoded by neuronal activity *per se*, but also by neuronal identity. This type of population code is referred to as labeled-line code (Aronov et al., 2003; Jia et al., 2010). Accordingly, higher-order sensory areas need to take into account *which* neurons are active when producing more specific representations of behaviorally relevant stimulus aspects (Quiroga et al., 2005).

Do the principles derived by optimal-coding theory also apply to networks with relatively few neurons and a restricted set of relevant stimuli? Here I investigate this question in the auditory periphery of grasshoppers. These insects produce genetically-fixed songs to recognize and evaluate potential mates with high fidelity (von Helversen, 1972). The involved processing stages comprise a feed-forward network of only three layers in the grasshopper's meta-thoracic ganglion: 60 receptors per side faithfully encode the signal's envelope (Machens et al., 2001; Rokem et al., 2006) and form the input stage; receptors project onto an intermediate layer of approximately 15 local neurons; these in turn connect to the output layer of ≈ 20 ascending neurons (Stumpner and Ronacher (1994); Vogel et al. (2005); Kumar et al. (2010), see Fig. 2.4). The output of this size-constrained network is the only source of acoustic information available to the behavioral decision centers in the brain.

What transformations does the neural representation of grasshopper song undergo in this small sensory system? Are these transformations similar to those found in larger sensory systems? I find that the neural representation changes profoundly across neuronal layers: sparseness and decorrelation of responses increases—just as in more complex systems and in accord with optimal coding theory. In the third layer, neuronal identity becomes crucial for an effective readout of the population. Within just two processing steps a labeled-line code is formed from a uniform representation of the stimulus at the input layer. This labeled-line code includes explicit representations of behaviorally-relevant stimulus features at a surprisingly early stage of the auditory pathway and presumably does not provide the complete representation of stimulus space found in the periphery of many larger sensory systems.

2.2. Methods

2.2.1. Recordings and Stimuli

Recordings were provided by Sandra Wohlgemuth and Olaf Kutzki. Single-unit intra-cellular recordings were performed from morphologically identified neurons at the 3 processing stages in the metathoracic ganglion of migratory locusts: receptors, local and ascending neurons. Each individual animal possesses the same set of roughly 35 unique, morphologically and physiologically identifiable types of local and ascending neurons. Cell types were identified morphologically by intracellularly injecting a fluorescent dye (Lucifer yellow). Recordings stem from 8 receptor neurons, 5 different types of local neurons (spanning the whole range of response types: TN1, BSN1, SN1, SN2, SN3) and 7 different types of ascending neurons (AN1, AN2, AN3, AN4, AN11, AN12, AN14)—most of them several times (range 1-10, median frequency 3). Each recording comes from a different animal. For details on the recording procedure and stimulus presentation refer to Wohlgemuth and Ronacher (2007).

Neurons were stimulated with 8 different calling songs of male grasshoppers of the species *Chorthippus biguttulus* used previously in Machens et al. (2003). As song periodicity (duration of syllable plus pause) depends on temperature in these poikilothermic animals (von Helversen, 1972; Creutzig et al., 2009), the stimuli were rescaled to a period of 100 ms and the carrier spectra were equalized (see Machens et al. (2003)). This leaves the stimuli differing only in their envelope's fine structure and probes the system at its limits of temporal resolution. Note, that I study responses of neurons of one species of acridid grasshoppers—*Locusta migratoria*—to courtship signals of another species—*Chorthippus biguttulus*. This is well justified as auditory neurons in the early stages of sensory processing are morphologically and physiologically highly similar (Ronacher and Stumpner, 1988; Sokoliuk et al., 1989; Neuhofer et al., 2008; Creutzig et al., 2009).

2.2.2. Estimation of response similarity, reproducibility, and sparseness

Time-varying firing-rate functions were estimated by binning spike trains at 0.05 ms resolution and smoothing them with a Gaussian filter of width $\sigma = 5$ ms. All results obtained were robust to changes of this filter's width.

Response similarity (cell-to-cell correlations) and reproducibility (trial-to-trial correlations) were quantified as Pearson's correlation coefficient. For response similarity trial-averaged rate functions of cell pairs for each stimulus were used. Reproducibility was based on pairs of single-trial responses to repeated representations of the same stimulus. Using the uncentered correlation coefficient as a measure of reproducibility gave similar results (Schreiber et al., 2003). Life-time sparseness of single neurons at each processing stage was quantified from their trial-averaged firing rate profiles using the quantity described in Willmore and Tolhurst (2001). Population sparseness was quantified from the trial-averaged firing rate profiles, by calculating the measure in Willmore and Tolhurst (2001) across the four cells in each population for every time bin and averaging over time. Each measure was averaged over all 8 songs for plotting and statistics. To not bias statistics in favor of cell types that

were recorded more often, I additionally averaged values over all specimen of a given cell type.

2.2.3. Spike-triggered averages

Spike-triggered average stimuli (STA) were estimated using the same responses as those used for decoding. To correct for different firing rates, all STA's were scaled to unit-norm. Individual STA's were averaged over specimen of the same cell type. Estimation errors were quantified by sub-sampling (10 repetitions using random subsets of 80% of the spikes). Estimation errors of the normalized STA's of neurons in all three layers did not differ significantly (mean$\pm$s.e.m. $9.2 \cdot 10^{-3}$, $p = 0.31$, one-way ANOVA).

2.2.4. Decoding

Information in neural responses was quantified using a decoding approach (Quiroga and Panzeri, 2009). To that end, song identity was decoded from neural responses using spike-train metrics (Victor and Purpura, 1997; van Rossum, 2001; Machens et al., 2003; Houghton and Sen, 2008). Although this approach underestimates the full information in the statistical sense, one probably comes closer to what a concrete, biologically plausible system can read-out from the spike trains studied here.

Single-neuron metric: The spike-train dissimilarity of single neurons was quantified using the van Rossum metric (van Rossum, 2001). Spike trains were binned with a resolution of 0.05 ms and filtered with an α-function: $\alpha(t) = \Theta(t)t\exp(-t/\tau)$, where $\Theta(t)$ is Heaviside's function. The parameter τ governs the temporal resolution of the metric. The Euclidean distance between all pairs of responses (8 repetitions of 8 song segments of different males, duration 200ms each) yields a distance matrix, which forms the base for the classification algorithm outlined below.

Multi-neuron metric: Population data was combined from single-cell recordings of four individual cells. This was justified, as neural activity in the network is entirely stimulus-driven. Hence, neurons are conditionally independent: there are no "noise" correlations between neurons, only signal correlations Brody (1999). Because I was interested in how the population code changed between processing stages, three different classes of 4-cell populations were created by combining different types of *either* receptors *or* local *or* ascending neurons. Thus, each population was characterized by a unique combination of 4 different cell types of a single layer. To not over-represent those populations which consist of cell types that were recorded more often, I averaged information rates and gains for each kind of population (that is combination of cell types) for plotting and statistics.

For a formal derivation of the multi-neuron metric see Houghton and Sen (2008). Application of this metric amounts to filtering the spike trains with an α-function, embedding the spike trains from multiple cells into a vector space, and then taking the Euclidean distance between different spike trains. The resulting distance matrix for each population is

then used to quantify stimulus discriminability through the classification algorithm. Thus, the only difference to the single-cell metric is that the spike trains of the cells comprising a population are embedded into a vector space.

The multi-neuron metric allows for different kinds of embedding, which is controlled by the "independence" parameter θ—the "angle between cells". This parameter allows to interpolate between two versions of a population code: a summed-population code and a labeled-line code. At $\theta = 0$ degrees, the metric corresponds to a summed-population code, where responses of different cells are embedded co-linearly. Information about which cell fired which spike is lost. This is optimal only if differences in the firing pattern between cells in a population are not relevant for the decoding task or if cells in a population are similarly tuned—this applies in this case to receptors and local neurons. In contrast, information about each spike's origin is fully retained in a labeled-line code which is implemented at $\theta = 90$ degrees (orthogonal embedding). This is desirable, if cells are tuned differently and represent different aspects of a stimulus, like the ascending neurons.

To illustrate that the labeled line decoder incorporates information about which neuron fired which spikes—the neuronal identity of spikes—I provide a simplified example of how three different stimuli can be distinguished with the summed-population and the label-line decoder, respectively, based on surrogate responses from two neurons. Figure 2.1 a shows the surrogate spike trains of both cells in response to the three different stimuli. To simplify the argument and without loss of generality, these spike trains were reduced to spike counts, which corresponds to applying a filter with a large time constant τ. In response to stimulus 1, cell A (green) and B (blue) fire 3 spikes each. Stimulus 2 evokes 1 spike in cell A and 5 spikes in cell B. The response pattern for stimulus 3 is inverted: now cell A fires 5 spikes and cell B only 1.

The summed-population decoder sums these spike counts prior to computing pair-wise distances between all stimuli. As the sum of spikes in cell A and cell B is the same, the population response to all three stimuli are represented by a 6; they cannot be distinguished. In contrast, the labeled-line decoder does not pool the two cells in the population. Here, each response is represented by an ordered pair of spike counts, which is different for each stimulus. This is also reflected in the resulting distance matrices (Fig. 2.1 b). As the summed population spike counts are the same for all three stimuli, the distance matrix has all zero entries and the summed-population decoder cannot discriminate between the three stimuli (information 0 bit). The labeled-line decoder, however, discriminates all three stimuli, as all off-diagonal entries in the distance matrix exhibit non-zero entries (information $log_2 3 = 1.6$ bit). The labeled-line decoder can distinguish stimulus one from both stimulus 2 and 3. In particular, it can also disambiguate stimuli 2 and 3, which differ only in neuronal identity of responses (both stimuli evoke one spike in one cell and 5 in the other, but in different order). This ordering is the major difference to the summed-population decoder and reflects the role of neuronal identity for the labeled-line decoder.

Classifier: Responses were classified using a nearest-neighbor clustering algorithm as in Machens et al. (2003). Nearness was given by the single or the multi-neuron metrics. One template spike train was randomly selected from each of the 8 songs. The remaining spike

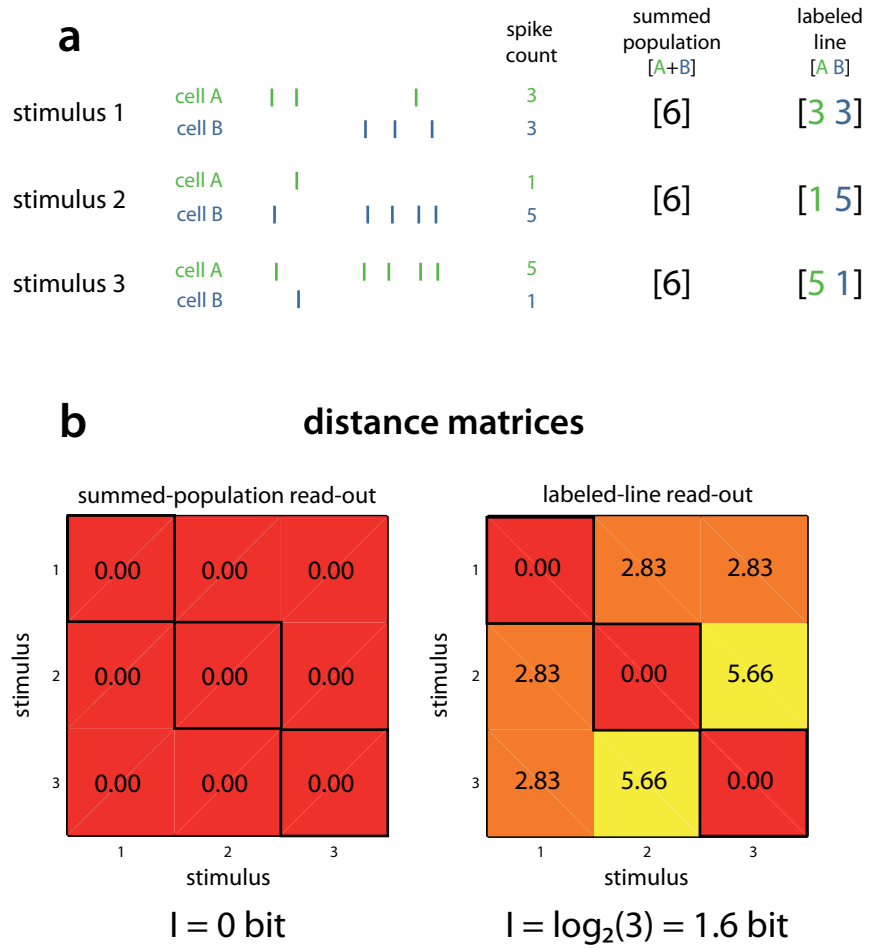

Figure 2.1.: **Illustration of the summed-population and the labeled-line decoders: a** artificially generated responses of two cells (A and B) to three arbitrary stimuli on the left. The right side shows the representation of the spike counts by the summed-population and labeled-line decoder. **b** Resulting distance matrices and information values.

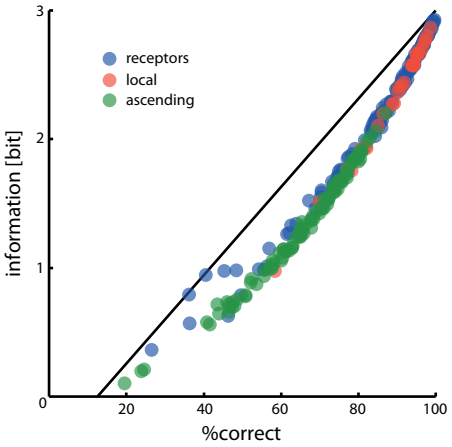

Figure 2.2.: **Two measures of decoder performance are highly correlated:** In the data set, both the mutual information and the percentage of correct classification yield highly similar results for decoding of single-cells as well as of populations ($r^2 = 0.97$).

trains were than classified as being evoked by the song the nearest template belonged to. This was repeated many times, always with a new, randomly selected set of templates. The classification results are organized in a confusion matrix $H(s, s')$, which shows the frequency with which a spike train being evoked by song s was classified as being evoked by song s'. The average of this matrix's main diagonal denotes the ratio of correctly classified spike trains.

Estimation of information: The mutual information of this confusion matrix $I(s, s')$ was used as a proxy for the information content of the neural responses $I(s, r)$ (Quiroga and Panzeri, 2009). Information is given by: $I(s, r) \propto I(s, s') = \sum_{s,s'} p(s, s') \log_2 \frac{p(s,s')}{p(s)p(s')}$ where $p(s, s')$ is the entry in the confusion matrix, $p(s) = \sum_{s'} p(s, s') = 1/8$ is the prior stimulus probability, and $p(s') = \sum_s p(s, s')$ is the marginal over the decoded stimuli (Victor and Purpura, 1997). Mutual information is 0 bit when the confusion matrix is "uniformly distributed", that is when each entry has the value $1/64$. It is maximal (for 8 stimuli $log_2(8) = 3$ bit) when there is a one-to-one relationship between spike trains and classes, e.g. when all entries are concentrated at the matrix' diagonal. As this measure is upwardly biased, I calculated the same quantity 10 times for random assignments between responses and stimulus classes and subtracted this bias from the naive estimator $I(s, s')$ (Aronov et al., 2003). Information was expressed either as a rate in bit/second by dividing the information by the stimulus length (maximal information rate being thereby 8/0.2 s = 15 bit/s) or as information per spike (bit/spike) by normalizing the information rate by the cell's firing rate. Firing rate was quantified as the spike count divided by the length of the spike train segment (200 ms).

Another common measure for classification success is the percentage of correct classification given by the sum over the diagonal of the confusion matrix. As a test, I compared this measure to the mutual information and found a high degree of correlation between both

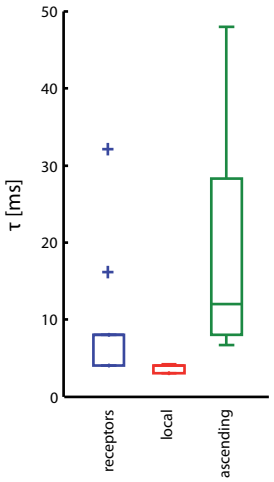

Figure 2.3.: **Optimal time scales for decoding:** Box plots show the τs that maximized the mutual information for each cell type. These determine the width of the α functions with which spike trains were convolved in the decoding procedure and indicate the time scale at which the decoder operated.

(Fig. 2.2).

Optimization of the metric's parameters: Classification performance is a function of the metric's temporal resolution τ. Information was optimized with a grid search for τ ranging from 0.25 to 64 ms (9 values, spaced linearly on a logarithmic scale). The τ used for decoding are shown in Fig. 2.3. Receptors exhibited an intermediate range of τ between 4 and 8 ms with two outliers at 16 and 32 ms. The τ of local neurons were significantly smaller ($p = 0.01$), spanning a range of 3 to 4.2 ms. Ascending neurons had the highest τ between 6.7 and 42 ms, being significantly greater than those of local neurons ($p = 0.003$). For population decoding with the multi-neuron metric, a single optimal τ for all cells in a population was used.

In the results section, I consider only the information rates obtained for two "extreme-value decoders" at $\theta = 0$ degrees (summed-population) and at $\theta = 90$ degrees (labeled line) for each population. I have also determined information at the *optimal* θ for each population by a grid search in the interval $[0, 90]$ degrees. As either of the two decoders at 0 or 90 degrees yielded near-optimal performance for any population (median information loss 2%), I decided to consider only those two for all analyses.

2.2.5. Statistics

All plots and statistics were based on average values for each cell type or type of population. That is, over all recordings of a cell type for the analysis of single cells and over unique, unordered 4-tuples for populations of cells. Tests—if not stated otherwise—were either parametric (t-test) or non-parametric (two-sided Wilcoxon's rank-sum test), depending on the outcome of a Jarque-Bera test for normality with a significance level $\alpha = 0.05$. No

correction for multiple comparisons was performed in order to avoid false negatives, as I was interested in the outcome of each individual pair-wise comparison, not in the general detection of statistical differences between groups.

All analysis was done in Matlab.

2.3. Results

Song recognition in grasshoppers is based on the signal's envelope. Carrier frequency is largely irrelevant, as the frequency resolution of the auditory system is weak and the carrier spectrum is relatively uniform across species Meyer and Elsner (1996, 1997); Hennig et al. (2004). The envelope of the grasshopper's calling song is composed of several repetitions of a basic subunit: a syllable-pause pair. I evaluated neuronal responses to two syllable-pause pairs, corresponding to the minimal signal duration necessary for song recognition by male grasshoppers (Ronacher and Hennig, 2004). Despite their high homogeneity, these songs are well discriminated at the level of auditory receptors (Machens et al., 2003) and in behavioral tests (Einhäupl et al., 2011).

2.3.1. Life-time sparseness increases and reproducibility decreases

First, I analyzed properties of single-neuron responses. Example spike trains of different receptors, local, and ascending neurons show that neuronal response characteristics changed within the network (Fig. 2.4). Receptors exhibited largely persistent firing. In contrast, local and ascending neurons appeared to respond more transiently with many interleaving periods of silence. Moreover, reproducibility of responses to the same stimulus tended to be lower for ascending neurons. To quantify this observation, I analyzed the life-time sparseness as well as the correlation-based reproducibility.

Life-time sparseness: It provides a measure of how much of a neuron's response is concentrated in a few, transient firing events and indicates a cell's selectivity for temporal features of the stimulus. It is not to be confounded with population sparseness, where few neurons are active at any time. Following the definition of Willmore and Tolhurst (2001), life-time or temporal sparseness is bounded between 0 (equal firing rate across the whole stimulus) and 1 (all firing concentrated at one point in time). While the average life-time sparseness of receptors was low (0.26 ± 0.15, mean $\pm$ std), local and ascending neurons fired more sparsely (0.52 ± 0.19 and 0.57 ± 0.12 respectively, $p < 0.01$, rank sum; Fig. 2.5 b). Quantification of temporal sparseness by the kurtosis of the firing-rate distribution yielded similar results (Fig. 2.6) Hence, temporal sparseness of song representation across the network is increased—a transformation suggested by optimal coding theory Olshausen and Field (2004); Smith and Lewicki (2006).

Reproducibility: Local neurons often exhibited transient and temporally-precise firing events. Although the ascending neurons showed a comparable degree of sparseness, their precision was lower and firing appeared to be more variable across trials. I quantified the reproducibility of responses by Pearson's coefficient of correlation between pairs of spike trains. Ascending neurons fired less reproducibly than receptors and local neurons (receptors 0.66 ± 0.13, local 0.68 ± 0.11, ascending 0.43 ± 0.13; $p < 0.019$ rank sum test; Fig. 2.5 c).

Overall, the neural representation of song was transformed in this three-layer network: temporal sparseness of the responses was established after just one synapse; reproducibility decreased at the third stage. This left the ascending neurons—the network's output layer—with a sparser yet more noisy representation of the song.

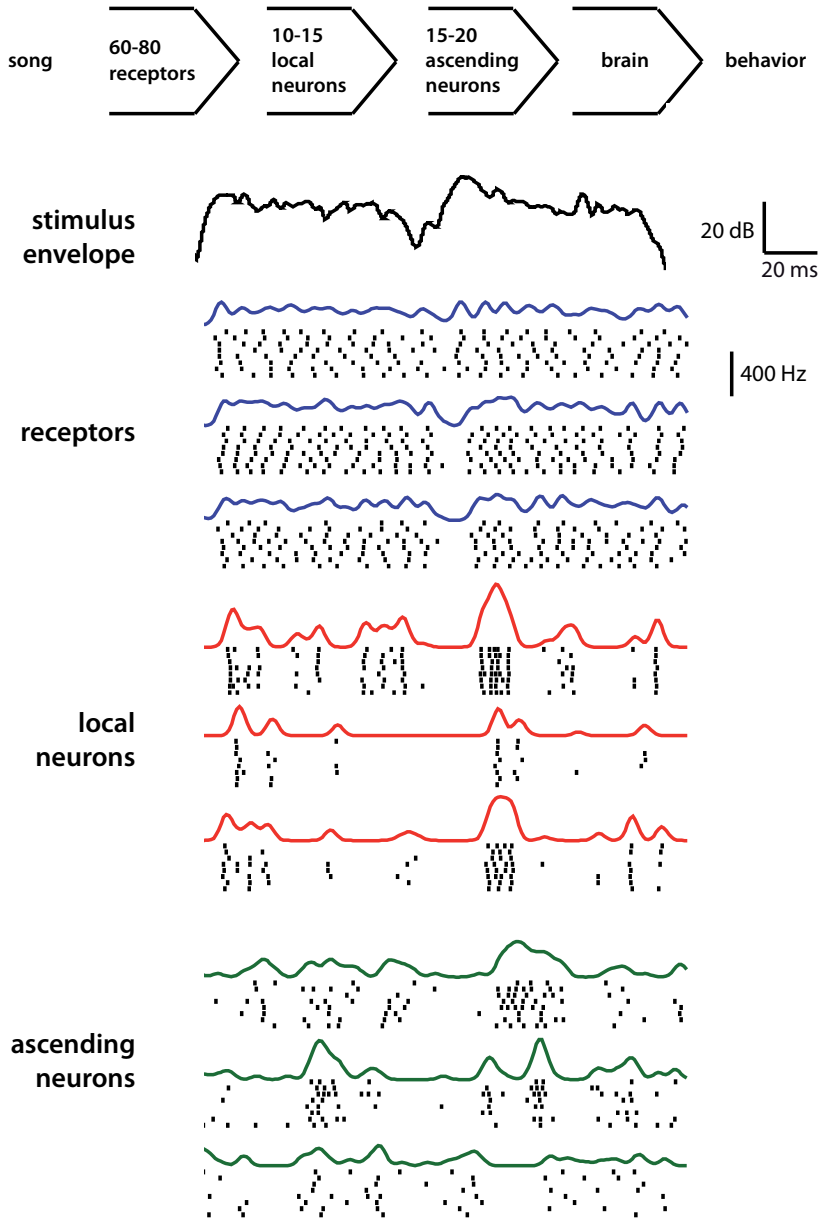

Figure 2.4.: The schema on top depicts the structure of the auditory system of grasshoppers. Below are responses of different receptors, and different types of local and ascending neurons (3 representatives each) to two syllables of the song of the grasshopper *Chorthippus biguttulus* (top black line). Short black lines mark the spike times for each of the 8 stimulus repetitions; thick colored lines above each block show the trial-averaged firing rate functions. Top scale indicates the amplitude and time scale for the song. Lower vertical scale marks the amplitude for trial-averaged firing rate functions; time scale of firing rates as in the song.

2.3.2. Information in individual neurons is reduced at the network's output layer

The observed loss in response reproducibility may compromise stimulus discriminability. On the other hand, increased sparseness of responses may increase robustness to noise and improve stimulus discriminability at the level of ascending neurons. To quantify the effect of those opposing trends, I analyzed the information transfer by decoding stimulus identity from single-neuron responses (Quiroga and Panzeri, 2009). Intuitively, stimulus discriminability—and hence information—is high, if the responses to different stimuli are less similar than those to repeated representations of the same stimulus. The decoding algorithm follows this intuition, by basing the classification of responses on the similarity of spike trains defined by their Euclidean distance (van Rossum, 2001; Machens et al., 2003).

Although the median information rate was higher in local neurons than in receptors, this increase could not be confirmed with high significance (Fig. 2.5 d, receptors 6.7 ± 3.8 bit/s, local neurons 8.7 ± 3.1 bit/s, $p = 0.33$, rank sum). Interestingly, the information rate of isolated cells decreased at the level of ascending neurons to 3.3 ± 1.9 bit/s ($p = 0.019$, rank sum), dropping even below that of single receptors ($p = 0.044$, rank sum). For comparison, I also show the classification success as quantified by the fraction of correctly classified spike trains (Fig. 2.5 f, receptors 61 ± 24 %, local neurons 76 ± 13 %, ascending neurons 43 ± 15

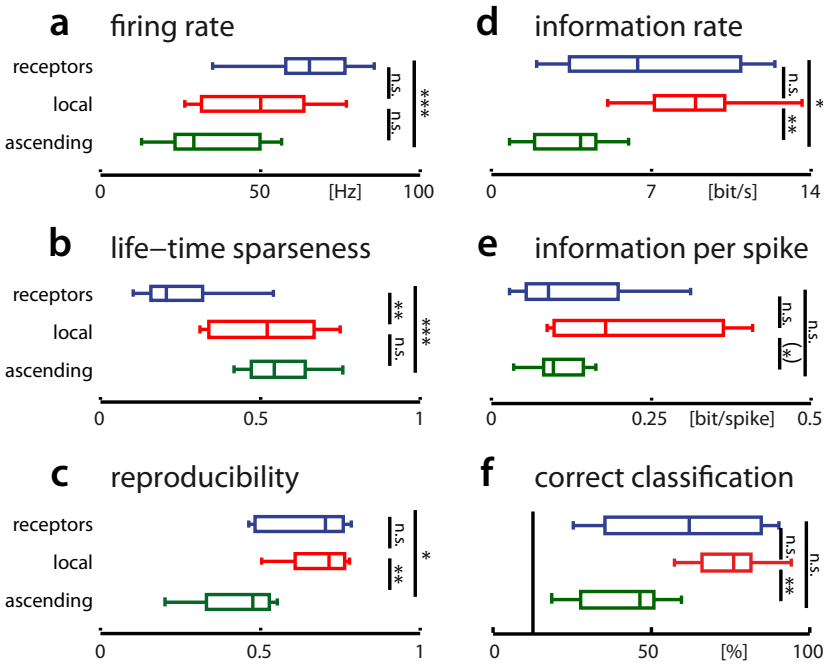

Figure 2.5.: **Response characteristics of individual neuron types in all three layers of the network: a** firing rate **b** life-time sparseness **c** reproducibility (Pearson's coefficient of correlation between firing rates of different trials) **d** information rate of single neurons **e** information per spike (information rate normalized by the firing rate of each cell); **f** percentage of correctly classified spike trains (the vertical line marks the chance level at 12.5%). n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Blue - receptors (N = 10 cells), red - local neurons (N = 21 cells of 5 types), green - ascending neurons (N = 25 cells of 7 types).

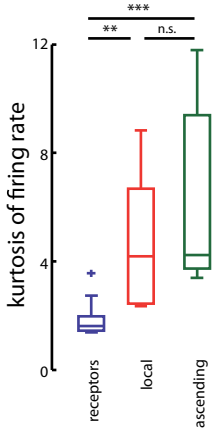

Figure 2.6.: **Kurtosis of the firing rate distribution** as an alternative measure of life-time sparseness. Receptors 1.9 ± 0.7, local neurons 4.8 ± 2.7, ascending neurons 6.2 ± 3.6; receptors vs local neurons $p = 0.008$, receptors vs ascending neurons $p = 2.1 \cdot 10^{-4}$, local neurons vs ascending neurons $p = 0.5$. n.s. $p > 0.05$, ** $p < 0.01$, *** $p < 0.001$, rank sum.

%). This measure of classification success was highly correlated with information ($r^2 = 0.97$, see Fig. 2.2), but yielded significant differences only between local and ascending neurons ($p = 0.003$, all other pairs $p > 0.09$, t-test).

Remarkably, information about stimulus identity retrievable from individual neurons is lowest at the output layer of the network. This is partly explained by their comparatively low firing rate (Fig. 2.5 a, e).

2.3.3. Ascending neurons decorrelate the neural representation of song

The higher reproducibility of neuronal responses in the first two layers (receptors and local neurons) would enable the animals to discriminate individual songs better than the noisy responses of ascending neurons in the output layer (Fig. 2.5 d, see also Rokem et al. (2006); Wohlgemuth and Ronacher (2007)). This may at first glance seem paradoxical. Obviously, the network is not enhancing the *raw* stimulus discriminability at the level of individual neurons. A likely explanation is that the population code for song undergoes a transformation at the level of ascending neurons: While local neurons might constitute a relatively homogeneous population—where every neuron encodes largely the same information—the ascending neurons possibly use a distributed representation, where different neurons encode different aspects of the stimulus. This might facilitate subsequent processing steps—at the expense of a lower *single*-cell information about song discriminability.

A comparison of the firing-rate functions of cells within a layer suggests that this is indeed true (Fig. 2.4): Responses to a given stimulus were diverse across different ascending neurons, while responses among the groups of receptor and local neurons, respectively, appeared to be more similar. I quantified the similarity of responses (same data sets as for the information estimation) by calculating Pearson's correlation coefficient between pairs of cells within each layer. The average response similarity steadily declined within the network

from 0.58 ± 0.14 in receptors over 0.42 ± 0.07 in local to 0.09 ± 0.16 in ascending neurons (Fig. 2.7 a, $p < 4.2 \cdot 10^{-4}$, rank sum). This decorrelation was accompanied by an increase of population sparseness across the network layers (Fig. 2.7 b, receptors 0.19 ± 0.04, local neurons 0.35 ± 0.05, ascending neurons 0.47 ± 0.03, $p < 5 \cdot 10^{-7}$, rank sum). Population sparseness measures to what extent only few neurons in a population are active at any time. This is not to be confused with temporal or life-time sparseness, which quantifies, how sparsely a single neuron fires over time (see Fig. 2.5 b). While the population sparseness of responses does not reach extreme values as e.g. in the olfactory system of locusts (see e.g. Laurent (2002), with values > 0.9), ascending neurons in the auditory system can still be considered to form a decorrelated and more sparse representation of the song (compare to Vinje and Gallant (2000)).

The decorrelation of responses in the output layer—i.e. the fact that different types of ascending neurons fired at different times during the stimulus—suggests that also feature selectivity may be more diverse. To show this directly, I calculated each neuron's spike-triggered average filter (STA)—the average stimulus preceding a spike—using the same set of spikes as before. Neurons at all three layers are well described by simple linear-nonlinear models (see e.g. Machens et al. (2001) and chapter 3). As the natural songs are strongly non-Gaussian (Machens et al., 2001), the STA does not represent an unbiased estimate of the neuron's filter (Sharpee et al., 2006). However, it allows one to assess the diversity of feature selectivity at each processing stage.

Figures 2.7 d–f show the STA filters of different receptors and different types of local and ascending neurons. Evidently, the variety of stimulus features eliciting spikes increased across the three neuronal layers: All receptors and local neurons had almost identical, unimodal STAs (Fig's 2.7 d and e, blue and red lines and Fig. 2.7 c, mean correlation between pairs of filters 0.91 ± 0.08 and 0.81 ± 0.09 respectively) and varied only little around the population average (Figs 2.7 d and e). In contrast, the filters of different types of ascending neurons were highly dissimilar: Some resembled those of the local neurons, while others exhibited strong negative components (Fig. 2.7 f, green lines, correlation between pairs of filters 0.22 ± 0.55).

This indicates that different types of ascending neurons encode different aspects of the stimulus. To further support my findings, I also looked at the similarity of each cell type's STA across different individuals/recordings, by computing Pearson's correlation coefficient between the average STA of a cell type and the STA of each individual cell of that type. In order for the STA (feature) to be specific for the cell type, this "intra-type" similarity should be larger than the similarity across different cell types of that layer ("inter-type" similarity).

As each receptor was considered a different type, inter- and intra-type similarities for receptors are identical (Fig. 2.8, blue box plots). For local neurons (Fig. 2.8, red box plots), both the STAs of the same cell type and of different cell types are highly similar (intra-type similarity 0.95 ± 0.04, mean$\pm$std, inter-type similarity 0.81 ± 0.10). While individual STAs of the same type are significantly more similar than those of different types ($p = 0.019$ rank sum), overall similarity at the level of local neurons is high. In ascending neurons, however, STAs of the same cell type are much more similar than those of different types (Fig. 2.8, green box plots, intra-type similarity 0.85 ± 0.17, inter-type similarity 0.22 ± 0.55, $p = 0.002$

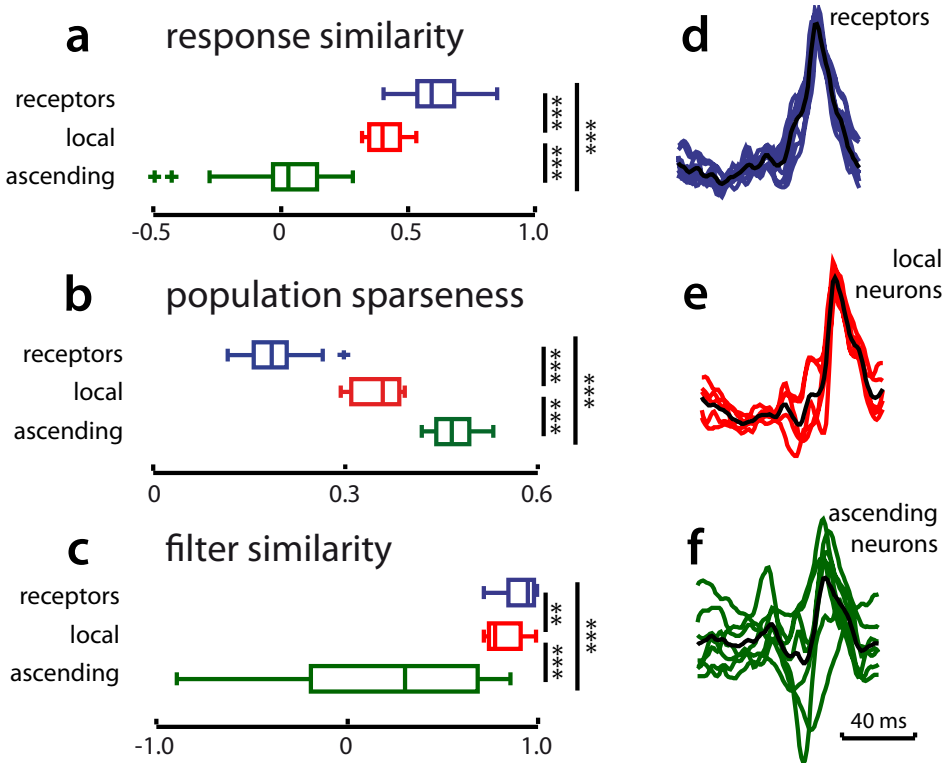

Figure 2.7.: **Population sparseness, diversity of response pattern and of feature selectivity increases for ascending neurons: a** Response similarity: correlation coefficients between firing-rate functions of pairs of neurons within each stage. **b** Population sparseness. **c** Filter similarity: correlation coefficients between the spike-triggered average filters of pairs of neurons within each stage. **d–f** Colored lines depict the normalized filters for individual cell types. The thick black line shows the average over all cell types in a population. Blue - receptors (N = 10 cells), red - local neurons (N = 21 cells of 5 types), green - ascending neurons (N = 25 cells of 7 types).

rank sum). This cell-specificity of STA filters further supports my hypothesis that each type of ascending neuron encodes a specific aspect of the stimulus.

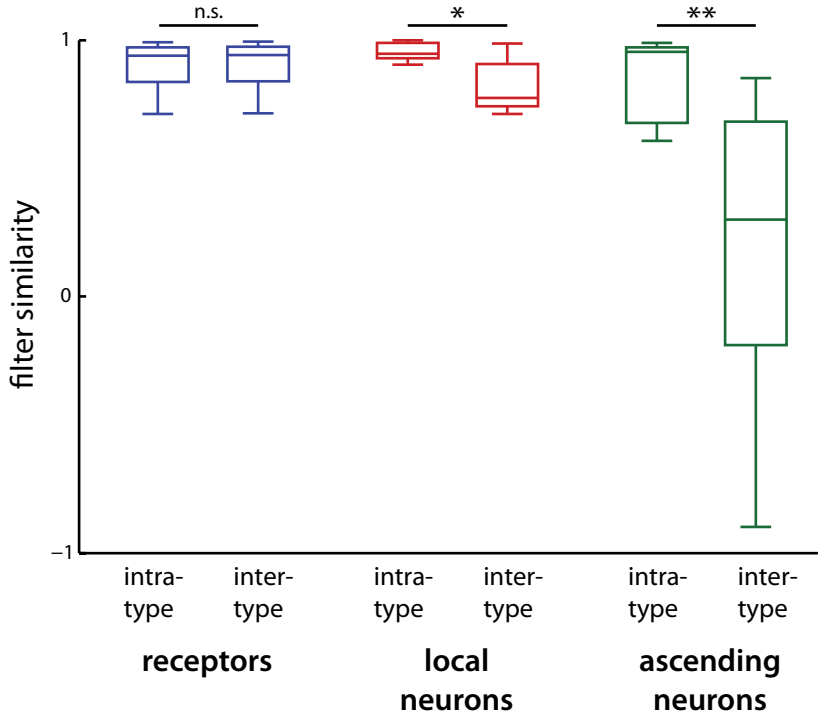

Figure 2.8.: **Cell-type specificity of STA filters:** Shown is the similarity of the STA filters of different specimen of the same cell-type (intra-type) and the similarity of the STA filters of different cell-types (inter-type, same as Fig. 2.7). n.s. $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, rank sum.

2.3.4. Ascending neurons profit most from a multi-neuron decoder

What is the consequence of the observed decorrelation at the network's output layer for the population code? In ascending neurons, information about song discriminability is likely to be distributed across different cells. As seen in the analysis of STAs (Fig. 2.7 f), individual neurons in this layer have a strong tendency to encode different aspects of the stimulus. In this case, blind pooling (summation) across neurons is likely to destroy valuable information, whereas knowing which cell fired which spike would be highly informative for a population read-out. Hence, a neuronal representation where not only the occurrence of spikes, but also their neuronal identity matters, is formed. Such a code corresponds to a *labeled-line* or distributed code (Aronov et al., 2003; Houghton and Sen, 2008).

In contrast, I expect neuronal identity to play only a minor role at the first two processing stages. Here, firing patterns and STAs within one stage (i.e. among receptors or among local neurons, respectively) are highly similar (Fig. 2.7d and e). Hence, pooling across neurons might increase information by enhancing the signal-to-noise ratio (see also Schneider and

Woolley (2010)). I refer to codes where neuronal identity does not contribute to an optimal readout as *summed-population* codes following Aronov et al. (2003) and Houghton and Sen (2008).

I tested how these two types of population code perform for the decoding of song identity in the three processing layers. To this end, I implemented two decoders on the basis of the multi-neuron metric proposed by Houghton and Sen (2008): one, which disregards information about neuronal identity of spikes by effectively summing up responses in a population—a "summed-population decoder"; and one, which preserves this information—a "labeled-line decoder", (see Houghton and Sen (2008), see Fig. 2.1). I constructed populations of four cells, consisting of four receptor, local, or ascending neurons, respectively (see Methods). For each of these populations, I decoded stimulus identity from responses using both decoders. First, I analyzed how much the readout of cell populations versus single cells can improve the discriminability of songs. Second, I determined which of the two population decoders was able to provide more information about stimulus discriminability in each layer.

Single-cell versus population decoding: Overall, information increased significantly for populations of receptors and ascending neurons when compared to single cells ($p < 7 \cdot 10^{-4}$, rank sum), but not for populations of local neurons ($p = 0.10$, rank sum; average information rate of populations of receptors 13.1 ± 1.2 bit/s, local neurons 13.0 ± 0.6 bit/s, ascending neurons 7.9 ± 1.6 bit/s). Evaluation of the performance of the population decoder in percent correct showed similar trends (receptors $93 \pm 5\%$, local neurons $94 \pm 2\%$, ascending neurons $72 \pm 8\%$).

While information increased for populations in all three processing stages, a true information gain in the population read-out compared to a single-cell read-out can only be expected, if the population decoder provides more information than the best single-cell within that population. I hence calculated the information ratio between the population as decoded by the better of the two multi-neuron metrics (determined individually for each population) and the best individual cell in this population. Gain values significantly greater than 1.0 signal a net increase of information when reading out the population compared to the best single cell.

Gain was intermediate for receptors (Fig. 2.9 a, 1.16 ± 0.11, $p < 4 \cdot 10^{-11}$ sign test against 1.0) and not significantly different from 1.0 for local neurons (1.05 ± 0.10, $p = 0.38$). Presumably, these two populations show very moderate gains because already the best single receptors or local neurons exhibited information rates close to the theoretical maximum of 15 bit/s and discriminated the stimulus almost optimally (see Fig. 2.5 d). This leaves little room for further improvement when considered as a population. Ascending neurons profited most from a population read-out. Here, information in the population increased on average 1.46 ± 0.21 fold compared to information in the best single cell ($p < 6 \cdot 10^{-11}$).

Alternatively, I also considered the gain with respect to the average information of all 4 cells comprising that population (Fig. 2.10). Clearly, this measure of information gain yields higher values: Receptors exhibit an average gain of 1.99, local neurons 1.49 and ascending neurons 2.49. Thus, the gain relative to the average information in the population is 1.4–1.7 fold greater than the gain relative to the best cell (compare Fig's 2.9 a and 2.10). This is due

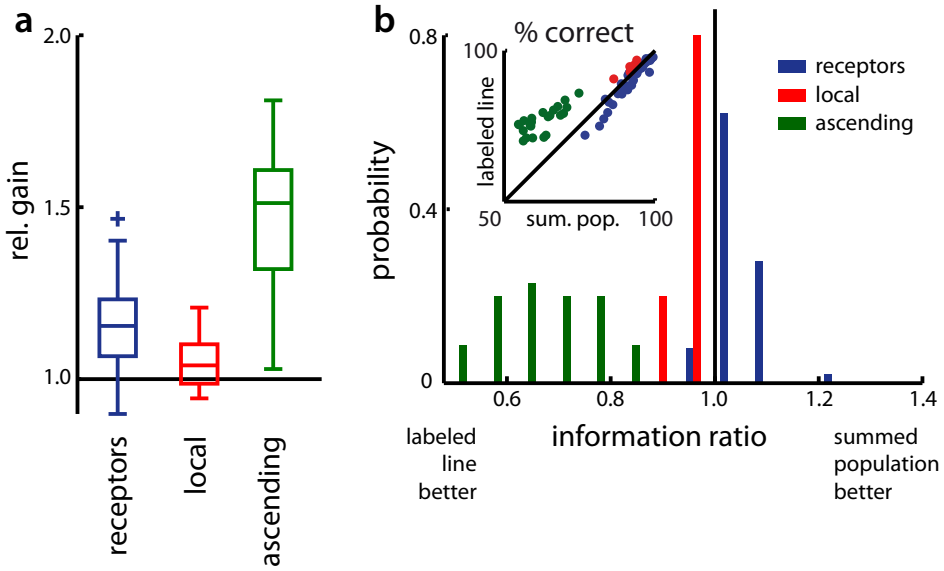

Figure 2.9.: **Population decoding: a** Relative gain of information through population decoding. **b** Optimal population decoder. Inset: Percentage of correctly classified responses for the summed-population and for the labeled-line decoder. The histogram shows the probability distributions of the ratio of information achieved by a summed-population and a labeled-line decoder for the three processing stages (thick black line at 1.0 corresponds to equality, values smaller than 1 indicate loss of information through summation). Blue - receptors (N = 100 four-cell populations), red - local neurons (N = 160 four-cell populations, 5 different combinations of cell types), green - ascending neurons (N = 910 four-cell populations, 35 different combinations of cell types); dots in inset correspond to averages over each combination of cell types.

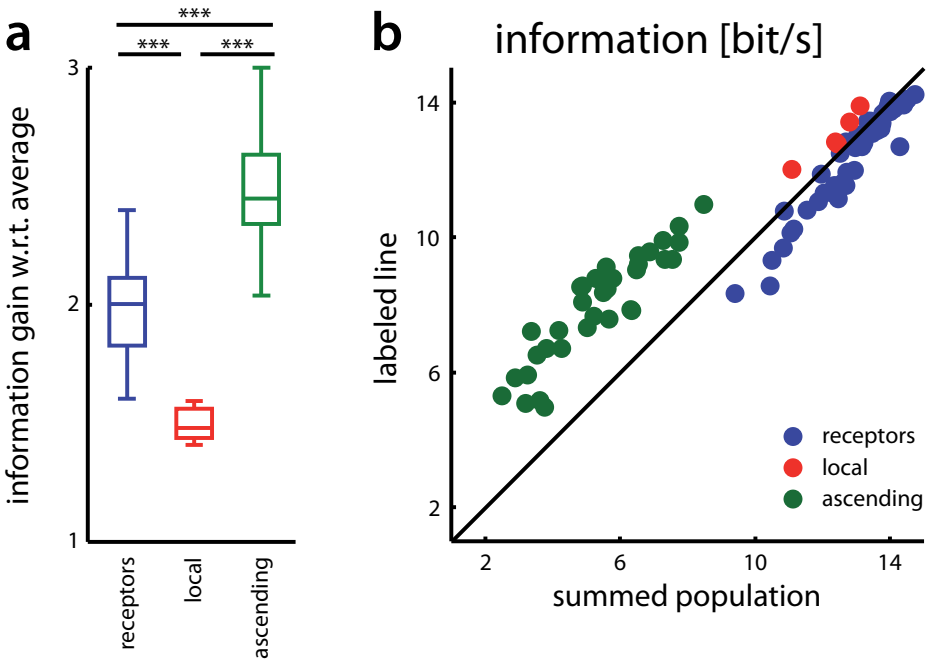

Figure 2.10.: **Information of the population decoder: a** Information gain of the population with respect to the average information obtained by decoding them individually. Receptors 2.0 ± 0.2, local neurons 1.5 ± 0.1 ascending neurons 2.5 ± 0.2, all $p < 6 \cdot 10^{-7}$, rank sum. Horizontal line at 1.0 indicates no gain. **b** Information of the summed-population vs. that of the labeled-line decoder. Results for decoder performance in terms of % correct look similar (see inset in Fig. 2.9 b). *** $p < 0.001$, rank sum. Blue - receptors (N = 100 four-cell populations), red - local neurons (N = 160 four-cell populations, 5 different combinations of cell types), green - ascending neurons (N = 910 four-cell populations, 35 different combinations of cell types).

to an upward bias in this alternative measure: The more cells one includes in a population, the more likely it is to "hit" a highly informative one. Especially the receptors with their high spread of single-neuron information rates (see Fig. 2.5 d) are susceptible to this bias. Hence, information gain relative to the *best* cell in each population is preferable as it is a more conservative and less biased measure.

Note that the network converges from the receptors to the local neurons (cell numbers across layers are reduced by a factor of approximately four). Because no information is lost between these layers, this processing step could serve to compress the code.

Optimal population decoding: The ratio of the information transfers obtained from the summed-population versus the labeled-line read-out indicates the costs of ignoring the neuronal identity of spikes in a population (Fig. 2.9 b, see also Fig. 2.10 b): If the ratio is close to 1 then no information is gained by considering the neuronal identity of spikes; a ratio < 1 indicates that the labeled-line decoder yields more information, showing that it is costly to ignore information about which cell fired which spike.

The distribution for receptors and for local neurons clustered around a ratio of 1.0—that is, they were read-out almost optimally using either decoder. However, there existed small yet significant trends in both cell groups: populations of receptors were read-out signifi-

cantly better as a summed population (median ratio 1.04, $p = 4 \cdot 10^{-8}$, sign test against 1.0); populations of local neurons were decoded significantly better as a labeled line (median ratio 0.94, $p = 2 \cdot 10^{-6}$, sign test against 1.0). These quantitatively weak effects in receptors and local neurons are in contrast to what I found in ascending neurons: Here, the ratio was on average 0.69 and hence much smaller than 1.0 ($p < 8 \cdot 10^{-28}$, sign test against 1.0). Hence, a large amount of information in the population would be lost by ignoring the neuronal identity in populations of ascending neurons. The multi-neuron decoder confirms my hypothesis that ascending neurons implement a labeled-line code, while neurons in the first two layers are read-out efficiently as a summed population.

This result is also in agreement with the observed change in single-neuron information between layers (Fig. 2.5 d): Single-neuron information increased from the first to the second layer. Considering that 60 receptors converge on 15 local neurons, noise is likely to be reduced at this stage by summation and consequently single-neuron information is enhanced. In the next step—the transition from local to ascending neurons—single-neuron information decreased. This is consistent with a labeled-line code, as information about song discriminability is distributed across the population; hence information in *individual* neurons is lower.

2.4. Discussion

I asked whether the principles derived by optimal-coding theory in the context of larger neuronal networks also apply to networks with relatively few neurons and a restricted set of relevant stimuli. Analyzing the representation of natural signals in the early auditory system of grasshoppers, I find that this small system performs transformations akin to those found in larger systems: both temporal and population sparseness of the neuronal representation increase and responses are decorrelated. A labeled-line code for temporal features of the grasshopper's song is formed within only two processing steps. This code, however, differs fundamentally from that of larger systems, both in terms of how it is created as well as in its degree of specialization.

2.4.1. The grasshopper labeled-line code is different from that of larger sensory systems

Usually, labeled-line representations in larger systems are either obtained by combination of inputs that were derived from a labeled line themselves or at the very periphery by specific tuning of receptor neurons, like location dependence in the retina or frequency selectivity in mammalian auditory receptor neurons. Preferred stimulus features in these labeled lines often vary systematically along an anatomical gradient following a topographic order. Auditory receptor neurons in grasshoppers, however, are relatively uniform in their selectivity for temporal features - in agreement with the finding that receptor neurons are best read out as a summed population. Hence, the labeled-line code observed at the third neuronal layer is not explicitly derived from another labeled-line code, but has to be established *de novo* by a transformation of the stimulus representation in the first two layers. The construction of this labeled line from uniformly tuned inputs is presumably achieved by adaptation and a well-timed interplay of excitation and inhibition (Hildebrandt et al., 2009; Ronacher and Stumpner, 1988). I will provide direct evidence for this in chapter 3, showing two kinds of computations in the auditory system of the grasshopper that contribute to sparseness and map to these two mechanisms.

Moreover, the grasshopper labeled-line code is less general. While labeled-line codes in early sensory systems of higher animals tend to provide a complete representation of stimulus space allowing for flexibility of later read-out, part of the grasshopper labeled line already explicitly encodes stimulus features of behavioral relevance. So far, three specific examples of direct extraction of relevant features by ascending neurons have been described. Recognition of the "right" male is extremely relevant for female grasshoppers, as mating with the "wrong" one—in terms of species or quality—would severely impact a female's reproductive success. One song parameter supporting species recognition is the pause length of the song's subunits. The ascending neuron AN12 has been found to encode this song feature in its spike count, thereby allowing the female grasshopper to recognize the species of the male by its song (Creutzig et al., 2009, 2010). Another aspect of a male, which is indicative of high quality, is the ability to avoid predators. A frequent consequence of a previous encounter with predators is the loss of a hind-leg. As males produce their calling song by rubbing both hind-legs over their wings, a "one-legged" male will produce a dis-

torted song with tiny gaps in the song's syllable. The ascending neuron AN4 is strongly inhibited by songs exhibiting such gaps. A strong co-tuning between the firing rate of AN4 and the behavioral response has been shown before (Ronacher and Stumpner, 1988), indicating that this neuron explicitly encodes a male's quality. A third song parameter strongly influencing female choice is the onset slope of the song's subunits. The spike count of the ascending neuron AN3 is strongly modulated by the onset slope of the pulses, rendering it a potential encoder of this feature (Krahe et al., 2002). Thus, properties of the song which strongly influence behavior (pauses between syllables, gaps within syllables, and syllable onset steepness, respectively) are encoded explicitly in the spike count of ascending neurons AN12, AN4, and AN3.

Such an early specialization at the auditory periphery seems efficient, as the number of available neurons as well as the set of relevant stimuli is restricted. Usually, complete representations are found in systems comprising many neurons, where also the range of potentially relevant stimuli is large and relevance is often acquired through attention or learning (Olshausen and Field, 2004; Leonardo, 2005; Vinje and Gallant, 2000; Zhaoping, 2002; Laurent, 2002). The auditory system of grasshoppers, however, has to discriminate only a restricted set of genetically-fixed signals.

Interestingly, the structure of the auditory network seems to be older than that of many songs it processes (Neuhofer et al., 2008). It might thus be *optimal for* the songs but it is certainly not *optimized* for them. The fact that one finds behaviorally-relevant information represented explicitly in the network is likely to be a consequence of evolutionary adaptation of the songs to the receiver's network and not vice versa (Clemens et al. (2010); Neuhofer et al. (2008), see also Smith and Lewicki (2006)).

2.4.2. Trading "when" for "what" facilitates the read-out of long communication signals

The peripheral auditory system of the grasshopper seems to fulfill two functions of optimal coding: signal compression and facilitated read-out (Olshausen and Field, 2004; Barlow, 2001). Compression is observed from the first to the second neuronal layer, where neuronal numbers converge by a factor of four and information rates per neuron increase. Facilitated read-out is established at the neuronal output layer, as outlined in the following.

Receptors and local neurons encode temporal features of the song in the temporal structure of their spiking responses. They have to employ a temporally-precise and reproducible spike code in order to represent fine temporal features of a song (Rokem et al., 2006; Vogel et al., 2005), even preserving information about when a temporal feature occurred during a song syllable. In the recoded labeled-line representation of ascending neurons, however, the presence of a highly specific feature can be represented by the spike count across a whole syllable, alleviating the need for temporally-precise responses at the cost of reduced information about a feature's exact timing within a syllable. In that sense, the system trades the "when" of temporal features for an easily-decodeable spike-count representation of their "what". The observation that ascending neurons are best read out at larger time scales than local neurons is consistent with this hypothesis (Wohlgemuth and Ronacher (2007) and Fig.

2.3). Behavioral experiments suggest that the grasshopper's brain indeed often evaluates spike count and makes little use of spike timing (Creutzig et al., 2010; von Helversen and von Helversen, 1998). In addition, the songs exhibit great temporal redundancy: each song consists of 10-30 syllables allowing for multiple "looks" at the basic subunit (Viemeister and Wakefield, 1991). The signal's temporal redundancy presumably compensates for the restricted options of neuronal redundancy in this size-limited system and agrees well with a spike-count code that enables the system to increase the signal-to-noise ratio by accumulation of spikes across syllables. In chapter 4, I will show that a large part of song recognition in the grasshopper can indeed by explained with a classifier that does not use the "when" of temporal features, but only their "what".

2.4.3. Conclusion

Despite its limited size, the auditory system of the grasshopper shares properties of larger sensory systems: a sparse and decorrelated representation of inputs including a labeled-line population code. In contrast to larger systems, however, part of the auditory pathway seems to specialize early-on for specific behaviorally-relevant stimulus features. This representation is more reminiscent of higher-order areas in vertebrates. It is likely to restrict the set of stimuli that can be differentiated and hence to lower the flexibility of behavioral responses. In addition, information about precise timing of an event seems to be sacrificed for a pure detection of this event within a larger temporal window. In the context of mate selection, where signals have evolved to be sufficiently long and redundant, this may be a price worth paying and help to invest the limited available resources specifically in the extraction of relevant information.

3. Nonlinear computations underlying temporal and population sparseness in the auditory system of the grasshopper

Having established that the early auditory system of grasshoppers creates a temporal and population sparse representation of song from dense and uniform inputs (chapter 2), I will now show computations underlying this transformation of the neural code. To that end, I will fit linear-nonlinear models to electrophysiological recordings and find the structural elements being responsible for sparse coding in the grasshoppers. Using prior knowledge about the system, I will propose putative biophysical mechanisms implementing these computations.

3.1. Introduction

Successful behavior is tied to the formation of specific representations of the environment to enable the correct discrimination of friend and foe. Typically, complex feature selectivity arises from more generic representations as one ascends a sensory pathway. This increase of specificity can lead to temporal and population sparseness (but see Willmore et al. (2011)). Temporally sparse responses are characterized by well isolated firing events interleaved by periods of neuronal quiescence. Additionally, higher-order neurons are more idiosyncratic, each responding to "its own" feature. Consequently, neurons in a population are less prone to fire together, yielding population sparse activity. Besides being an outcome of the creation of more specific representations, sparseness also has intrinsic advantages like the efficient use of energy and neuronal bandwidth as well as the facilitation of subsequent computations involving learning and memory (Barlow, 2001; Olshausen and Field, 2004).

Here, I provide insight into how temporal and population sparseness arise in the auditory system of the grasshopper, which generates a sparse and specific representation of courtship signals in a small, three-layer feed-forward network as shown in the previous chapter.

Primary auditory receptors in the grasshopper yield a temporally dense, relatively unspecific and faithful representation of a sound's envelope Machens et al. (2001); Gollisch et al. (2002); Rokem et al. (2006). This code is then transformed to temporally and population sparse one second- and third-order neurons—the local and ascending neurons. In this chapter, the computations underlying this transformation will be characterized. In local neurons, a temporally sparse code is established. However population sparseness is low, as different local neurons respond to very similar features. Then, the ascending neurons generate a population sparse representation, where different ascending neurons respond to different stimulus patterns.

I fit low-dimensional models to recordings of second- and third-order neurons in the grasshopper using the framework of linear-nonlinear (LN) models. These models provide intuitive phenomenological depictions of the neural computations actualized by a neuron on its inputs. In their simplest form, LN models consist of a single linear filter and a static nonlinearity. As the transformation at the level of ascending neurons suggests more complex transformations, I used an extension of the one-filter LN models that allows one to describe multidimensional computations—spike-triggered covariance analysis (see e.g. Fairhall et al. (2006); Rust et al. (2005); Petersen et al. (2008); Fox et al. (2010).

I found two different classes of computation contributing to sparseness in the auditory system of the grasshopper: sensitivity to the derivative of a stimulus and a highly nonlinear, `and-not` like transformation. These abstract computations can be implemented by mechanisms ubiquitous in neural systems and are thus likely to constitute general mechanisms providing sparse and specific representations.

3.2. Methods

Animals, electrophysiology and acoustic stimulation Recordings were performed in adult locusts (*Locusta migratoria*) obtained from a local supplier and held at room temperature (22 ± 5 degree Celsius). I recorded intracellularly from identified auditory neurons in the locust's metathoracic ganglion. Auditory neurons are organized in a three-layer feed forward network with receptors as an input layer, an intermediate layer of local neurons and an output layer of ascending neurons. The data set consists of 8 types of auditory neurons from the intermediate (second-order or local neurons: TN1 N = 10, SN1 N = 2, SN3 N = 1, BSN1 N = 9) and the output layer (third-order or ascending neurons AN1 N = 5, AN2 N = 1, AN3 N = 2). Previous studies have shown that the local neuron BSN1 comes in two subtypes, one responding with a short burst to the onset of pulses and one firing more persistently during a pulse, most likely due to different strengths of inhibitory inputs (Stumpner and Ronacher, 1991; Stumpner, 1989). Accordingly, I refer to them as "phasic" (N = 6) or "tonic" (N = 3) subtypes of BSN1. Intracellular electrophysiological recording methods are described in detail in Vogel et al. (2005). After completion of the stimulation protocol, neurons were stained with Lucifer Yellow and identified by their characteristic morphology (Römer and Marquart, 1984; Stumpner and Ronacher, 1991).

Natural songs of grasshoppers consist of a broad-band carrier whose amplitude is modulated by a species-specific envelope. As the decisive cues for song recognition lie in this envelope, I were interested in how single neurons represent the pattern of amplitude modulation of a sound. The amplitude of broad-band noise ($5 - 40$ kHz) was therefore modulated with low-pass Gaussian noise (cutoff frequency 140 Hz). The mean of this amplitude modulation was set to $\approx 10 - 15$ dB above each cell's threshold (thresholds ranged between 45 and 65 dB SPL). The standard deviation of the random amplitude modulations was 6 dB. These noise stimuli were presented in two variants to estimate and verify the models: one long segment lasting between 5 and 14 min for estimating the models and a shorter 6 s segment which was repeated at least 18 times and was used for estimating the time-varying firing rate for model testing. For all further analysis, only steady-state responses were used, by omitting the first 400 ms of each spike train.

Describing neural computations using lower-dimensional models Responses to the long noise stimulus formed the basis for spike-triggered analysis (Schwartz et al., 2006). In essence, spike-triggered analysis consists of finding stimulus features influencing a neuron's spiking by comparing the distribution of stimuli $\vec{s}$ preceding a spike r, $p(\vec{s}|r)$, to the distribution of all stimuli, $p(\vec{s})$, and finding directions in stimulus space for which both distributions differ most. This yields linear-nonlinear cascade models of neural computation: a high-dimensional stimulus is reduced by linear projection to one or two features; then, a nonlinearity transforms the feature value(s) into the cell's firing rate.

Different versions of spike-triggered analysis differ in the divergence measure used to quantify the difference between $p(\vec{s})$ and $p(\vec{s}|r)$: In its simplest form, the difference of first and second central moments is maximized, yielding the spike-triggered average and spike-triggered covariance, respectively (Schwartz et al., 2006). Other approaches use the

Kullbach-Leibler distance as a divergence measure (Sharpee et al., 2004; Pillow and Simoncelli, 2006).

The stimulus $\vec{s}$ was defined as a vector corresponding to the envelope of the sound in the 64, 1 ms wide bins preceding each point in time. The 64-dimensional distribution of stimuli $p(\vec{s})$ was by construction Gaussian. $p(\vec{s}|r)$ was sampled by the spike-triggered ensemble (STE), that is, the set of stimulus segments preceding each spike collected in response to the long noise segment.

In its simplest form, spike-triggered analysis results in calculating the difference of the mean of both distributions, yielding the spike-triggered average (STA) as a single feature: $\vec{f}_{STA} = \sum_{\vec{s}} p(\vec{s}|r)\vec{s} - \sum_{\vec{s}} p(\vec{s})\vec{s}$ (the last term is the mean of all stimuli and a constant for the noise stimuli).

To characterize more complex, multi-dimensional feature selectivity, spike-triggered covariance analysis (STC) was performed. To that end, the difference ΔC of the covariance matrix of the STE $C_{p(\vec{s}|r)}$ and the covariance matrix of all stimuli $C_{p(\vec{s})}$ was computed: $\Delta C = C_{p(\vec{s}|r)} - C_{p(\vec{s})}$. The covariance matrix C of an arbitrary distribution $p(\vec{x})$ is given by $C_{p(\vec{x})} = \sum_{\vec{x}} p(\vec{x})(\vec{x} - \langle\vec{x}\rangle)(\vec{x} - \langle\vec{x}\rangle)^T$, where the angled brackets denote the average. An eigenvalue decomposition of ΔC yields stimulus directions in which the variance—and not the mean as for the STA—of the spike-triggered and the raw stimulus ensemble differ most. These directions are indicated by eigenvectors associated with non-zero eigenvalues. However, due to the finite sample size (number of spikes), most eigenvalues are non-zero. The significance of the deviation of each eigenvalue from zero was checked by computing 1 % and 99 % confidence intervals for the maximal/minimal eigenvalues of each recording: I generated randomized responses by shuffling the spike times and used the distribution of the larges/smallest eigenvalue from 1000 such randomized responses to derive confidence intervals. All cells in the data set exhibited at least two significant eigenvalues at this 1 % significance level.

The STC analysis was performed in a subspace orthogonal to the STA by projecting the STA from each stimulus vector: $\vec{s}^{\,o} = \vec{s} - (\vec{s}^T \vec{f}_{STA})\vec{f}_{STA}/|\vec{f}_{STA}|^2$. This rendered the STC eigenvectors orthogonal to the STA and often improved model performance. As eigenvectors yielding the filters recovered by STC analysis are only defined up to an arbitrary sign, the sign was chosen such that the STC filter is most similar to the negative derivative of the STA filter (Fairhall et al., 2006). Furthermore, all filters were normalized to unit-norm.

The nonlinearity is given by Bayes' rule as the ratio of the raw and the spike-triggered stimulus distribution in the stimulus subspace defined by the filters $\langle r\rangle \cdot p(\vec{s}'|r)/p(\vec{s}')$. $\langle r\rangle$ is the average firing rate in the response set used for estimating the model. $\vec{s}'$ is the stimulus projected onto a subspace defined by the STA, or the STA and the STC filter with the largest absolute non-zero eigenvalue; it can thus be either one or two-dimensional. $p(\vec{s}')$ is the distribution of projection values of all stimuli and is by definition Gaussian with standard deviation 6 dB. $p(\vec{s}'|r)$ is the distribution of projection values of the STE and was estimated by kernel-density estimation.

Two kinds of model were constructed for each recording: One model consisted only of the STA and a one-dimensional nonlinearity. It is referred to as the "STA model". The other model contained the STA filter and the filter with the largest absolute eigenvalue from the

STC analysis—here called the “STC filter"—plus a two-dimensional nonlinearity. It is called the “STC model".

Data analysis – quantification of model performance A bias-corrected version of Pearson's coefficient of correlation ρ was used to quantify how well each model predicted the neuronal response to a novel stimulus (Petersen et al., 2008). To that end, the time-varying firing rate $r(t)$ of the neuron was estimated from responses to several repetitions of a stimulus not used for model estimation by binning time at 1 ms and smoothing with a box kernel spanning two bins. A predicted response $\hat{r}(t)$ to the same stimulus was obtained from the STA and STC models.

As the neuronal response is noisy, a model of the neuron can never perform better than that noise level. Thus, the naive estimator of the correlation is downwardly biased by that noise. In order to correct for this bias, the noise in the response was estimated by calculating $r(t)$ from two equal-sized, exclusive subsets of the stimulus repetitions, yielding two independent rate-estimates $r_1(t)$ and $r_2(t)$. The coefficient of correlation between these two estimates was then used to normalize the raw correlation: $\rho' = \rho(r, \hat{r})/\rho(r_1, r_2)$.

Data analysis – characterization of model structure To characterize the shapes of the filters two metrics were used: the first was defined by the coefficient of correlation between the derivative of the STA filter and the STC filter. The second was given by the delay between the peak of the STA filter and the peak of the STC filter.

To characterize the kinds of computations described by the two-dimensional nonlinearity, I quantified which of the four quadrants drove the cell most. The STC nonlinearity was divided into four quadrants, defined by the two principal axes. The relative weight each quadrant had in driving the cell was computed by summing over all values in a given quadrant and normalizing by the sum over all quadrants. To quantify the contribution of positive or negative projections of the STA to firing, I calculated the relative weights of the two left and two right quadrants. For the STC filter, the relative weights of the upper and lower right quadrants, corresponding to positive projection values of the STA and negative/positive projections values of the STC filter were calculated.

Additionally, I asked what logical operations on the output of the STA and STC filters the two-dimensional nonlinearities implemented. To that end, the correlation between the nonlinearity and a template which drives firing only in selected quadrants was calculated. As templates, three two-dimensional nonlinearities were used: One which drives the cell only for positive output values of the STA filter ("STA"), one which drives it only for positive outputs of the STA and the STC filter ("STA `and` STC"), and one which drives it only for positive outputs of the STA and negative outputs of the STC filter ("STA `and not` STC"). The template was normalized such that the sum over all entries was 1.0. Hence, the correlation could range between 0.0 (no match) and 1.0 (perfect match).

Simulation of model responses to natural songs To study the responses of the models to natural signals, a set of songs from 8 different male grasshoppers of the species *Chorthippus biguttulus* was used (see methods in chapter 2). The natural songs had a standard deviation

of 6 ± 1 dB, close to the standard deviation of the noise stimuli used for the estimation and evaluation of the models. In order to cover the range of firing rates between 20 and 50 Hz observed for natural stimuli (see Fig. 2.5 a in chapter 2) the average amplitude was set to +6 dB.

Quantification of temporal and population sparseness Sparseness of the modeled responses was quantified using the same measures as in chapter 2 (Willmore and Tolhurst, 2001). For population sparseness, I constructed 500 populations for each model class by randomly combining the responses of 4 cells of the same model class. This particular population size was chosen to enable comparison with the measured values of populations sparseness in chapter 2 (Fig. 2.7).

3.3. Results

In the following, I will start by describing the models of one representative cell in detail. I will then show the two classes of computation I found in the data set and relate their properties to the generation of a sparse representation of natural communication signals.

3.3.1. Two-dimensional models capture additional aspects of computation

To provide an intuition for the model structure obtained by spike-triggered covariance analysis, I will describe the model estimated for one ascending neuron AN1 in depth. The one-dimensional STA model (Fig. 3.1 a, b) consisted of a single spike-triggered average (STA) filter, which describes the temporal feature the cell is responsive to, and a one-dimensional nonlinearity, which transforms the output of the filter to the cell's firing rate and depicts the neuron's tuning for that feature. The cell's STA was largely uni-modal, exhibiting one prominent positive lobe at -20 ms preceding the spike and a weak negative lobe between -30 and -50 ms preceding the spike (Fig. 3.1 a). Thus, the cell was mainly sensitive to a low-pass filtered version of the amplitude of the stimulus. The nonlinearity was skewed towards positive filter values, indicating that the cell preferred stimuli that were similar to the STA (Fig. 3.1 b). Firing was reduced for very large projection values—the cell exhibited thus a band-pass like tuning for the STA.

The spike-triggered covariance (STC) model (Fig.3.1 c, d) consisted of the STA and a second filter recovered by STC analysis. This STC filter was broader than the STA filter, mainly positive and led the STA (Fig. 3.1 c). Thus, the neuron was influenced not only by the STA but also by a low-pass filtered version of the envelope, 20 ms preceding the STA. The nonlinearity depicts how the two filters interacted to govern the cell's firing: The stimulus was filtered by both the STA and the STC filter in parallel, producing two values per stimulus; accordingly, the nonlinearity, was two-dimensional, returning a firing rate for every pair of filter values (Fig. 3.1 d). The nonlinearity associated with both filters, showed that the cell was best driven by stimuli in the lower right quadrant. That is, stimuli exhibiting positive projection values onto the STA and negative projection values onto the STC filter evoked the highest firing rates. This corresponds to an `and-not` like logical operation on the output of the filters. The STC filter suppressed firing, as the cell fired strongly only for negative projection values of it.

Thus, the STC model yielded a much richer description of neuronal feature selectivity of this cell: addition of a second filter revealed a highly nonlinear computation performed on the stimulus which was not obvious from the STA model alone. Both the STA and the STC model predicted responses to novel stimuli well (Fig. 3.1 e). The more complex STC model yielded slightly higher performance values than the STA model (0.57 and 0.63, respectively). This means that the two-dimensional model explained approximately 9 % more of the variance in the responses than the STA model.

Generally, the stimulus transformations of auditory neurons in the grasshopper were well described by the STA and the STC model—model performance ranged between 0.5 and 0.8. The two-dimensional STC models were able to capture additional aspects of the stimulus-response relation as they performed significantly better than the one-dimensional STA mod-

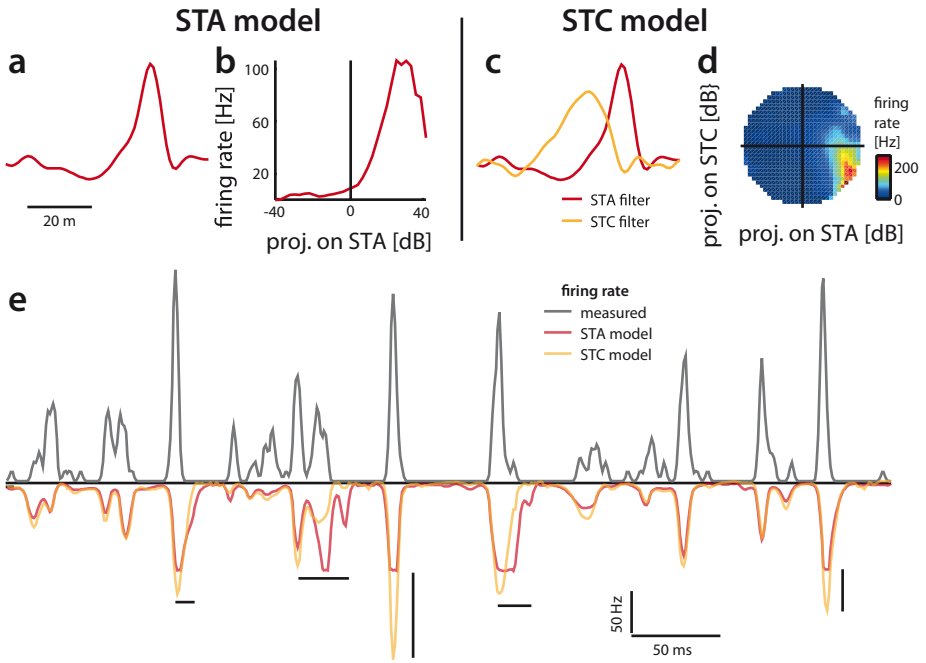

Figure 3.1.: **Structure of STA and STC model for an ascending neuron AN1.** **a** Spike-triggered average filter (STA). **b** Nonlinearity relating the projection values of each stimulus onto the STA to the cell's firing rate. **c** STA filter (red) and filter obtained by spike-triggered covariance analysis (STC, orange). **d** Two-dimensional nonlinearity relating the projection values of spike-triggered stimuli onto the STA (x-axis) and STC filter (y-axis) to the cell's firing rate. Firing rate is color coded (see colorbar). Note, that the neuron responds only to a small subset of projection values. **e** Measured response (black) and the predictions from the STA (red) and STC model (orange) were highly similar. The more complex STC model excelled during episodes marked by strong (vertical black bars) or prolonged firing (horizontal black bars).

els increasing model performance on average by 9 % (mean±std $\rho_{STA} = 0.59 \pm 0.11$, $\rho_{STC} = 0.65 \pm 0.11$, $p = 6 \cdot 10^{-6}$, sign rank).

However, there existed systematic differences in the structure of the predicted responses: First, the STC model in the example reached higher maximal firing rates than the STA model (Fig. 3.1 e, black vertical bars), due to the 2D nonlinearity covering a larger dynamic range—the 1D nonlinearity peaked at 118 Hz while the STC nonlinearity reached values up to 250 Hz (compare Fig's 3.1 b and d). Second, the firing epochs of the STC model in the example tended to be less sustained (Fig. 3.1 e, black horizontal bars). As we will see below, these differences enabled the two-dimensional STC models to exhibit increased temporal sparseness.

3.3.2. The model structure reveals two basic types of computation in the data set

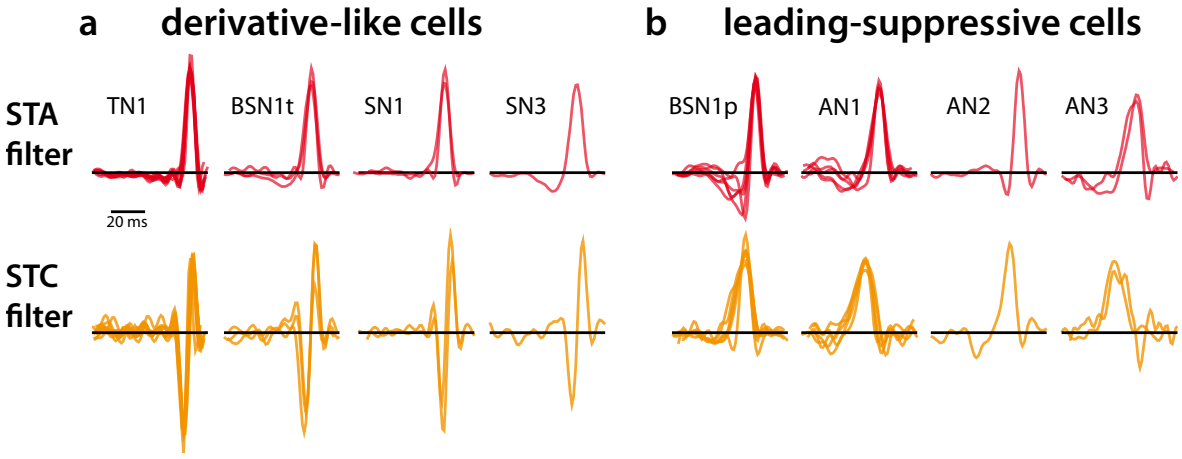

Figure 3.2.: **The cell types fall into two classes of model:** STA (top row) and STC filters (bottom row) for the cell types in the data set. Filters had a duration of 64 ms, were aligned at their peaks and normalized for better visibility. Only the STC filters differed strongly between classes while the STA filters of all cells were highly similar.

Looking at the model structure of all cells in the data set, I found two principal classes of models (Fig. 3.2). Notably, this dichotomy was not obvious by looking at the STA filters alone: the STA filter and its nonlinearity were similar to that shown in the example (Fig. 3.1)—the STA filter of all cells was thus mainly integrating and drove the cells for positive projection values (Fig's 3.2 upper row and 3.3 g, i). Only the incorporation of the STC filter yielded models with fundamental, qualitative differences, justifying the discrimination of two principal classes of neurons: the STC filters on the left side of the figure were all biphasic whereas those on the right were unimodal (compare Fig. 3.2 a and b, lower row). The fact that I found only two principal classes does not exclude the existence of additional classes in the auditory system of the grasshopper—local or ascending neurons not in the data set might implement different kinds of computations.

Analyzing the filters and nonlinearities allowed me to interpret the computations performed by both model classes:

a. "derivative-like" cells: I found specimen of this first group of cells only among the second-order, local neurons: TN1, SN1, SN3 and BSN1 (the tonic subtype, termed BSN1t) (Fig. 3.2 a). As for all cells, the STA filter was excitatory for positive projections of the stimulus (Fig. 3.3 i). The STC filter of this class of models was highly similar to the negative derivative of the STA (Fig. 3.3 e, correlation coefficient between the derivative of the STA and the STC filter 0.83 ± 0.11). As one filter was the derivative of the other, the shape of the STA filter strongly determined that of the STC filter. As the STA filters of different cells belonging to this class were relatively similar, the STC filter did not increase the diversity of the temporal selectivity of these cells. In addition, both filters were strongly overlapping in time, making these cells respond to the stimulus on a short time scale of the order of the STA filter's duration (Fig. 3.3 f, delay between peaks -3.0 ± 0.6 ms). The nonlinearity of the second filter was quadratic, rendering them weakly phase invariant (Fig. 3.3 j). However, there existed a clear bias towards positive projection values (Fig. 3.3 h). These cells thus approximated an `AND`-like operation on the output of both filters (Fig. 3.3 h). As the STC filter resembled an upstroke of the stimulus (Fig. 3.2 a, lower row), the neurons fired most strongly during loud onsets in the stimulus (Fig. 3.4 c). This accentuation of onset responses increased the transience of responses and thereby likely contributed to temporal sparseness. Derivative-like cells encoded a combination of the intensity—by means of the STA filter—and the derivative—by means of the STC filter—of a sound's envelope.

b. "leading-suppressive" cells: The local neuron BSN1 (the phasic subtype BSN1p), and the three ascending neurons AN1 (see Fig. 3.1), AN2 and AN3 formed the second class of models, which was thus dominated by ascending neurons. In contrast to the derivative-like cells, where the STC filter is constrained to be the derivative of the STA filter, here, both filters were largely independent and covered a longer segment of the stimulus: The STC filter was mostly integrating and led STA by 7.8 ± 1.9 ms; both filters spanned between 30 and 40 ms of the stimulus (Fig's 3.2, 3.3). This class of cells implemented thus much more complex computations than the derivative-like cells. Along this line, the range of delays between both filters (4 – 11 ms, 3.3 f) equipped these cells with a much more diverse temporal selectivity than the comparatively uniform derivative-like cells and potentially contributes to population sparseness. The nonlinearity of the STA filter was excitatory (Fig. 3.3 i). In contrast, that of the second feature was symmetric and suppressive, that is, it suppressed firing for large output values of the STC filter (Fig. 3.3 j). The joint nonlinearity of both features approximated an `and-not` like computation (Fig. 3.3 d, h): these cells were driven most strongly only when STA, but not when the STC filter was activated. The optimal stimulus corresponded thus to a loud segment in the stimulus being preceded by a softer segment. Loud and constant stimuli activated both filters which reduced tonic firing. This likely increased the transience of responses and hence temporal sparseness. Due to the second filter being suppressive and leading, I termed this model class "leading-suppressive cells".

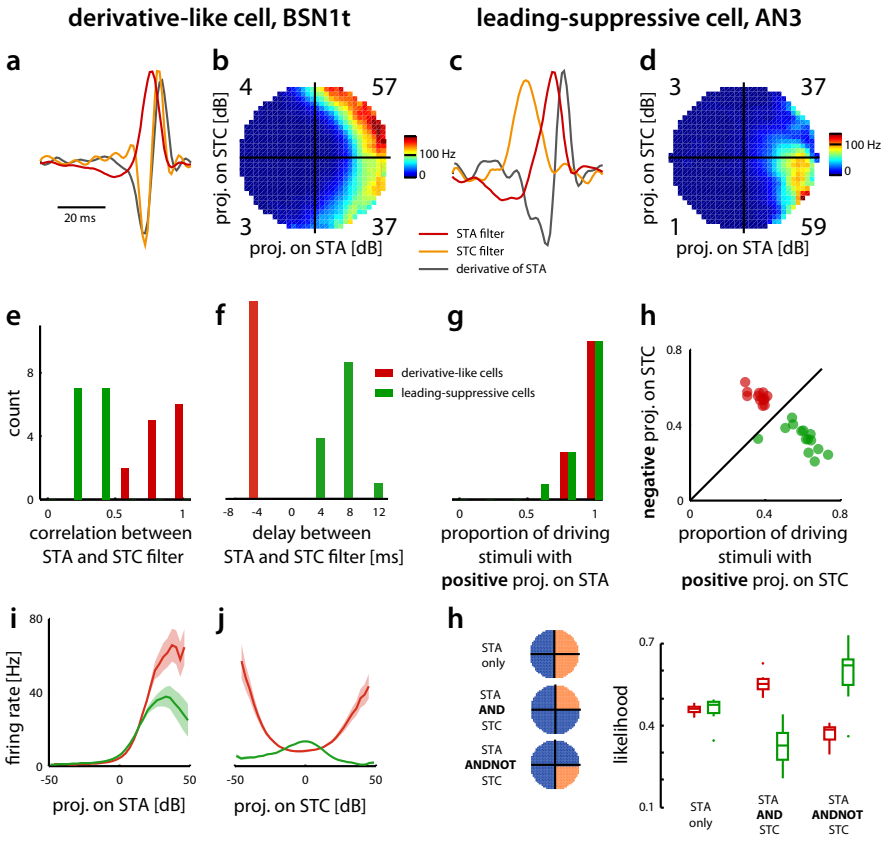

Figure 3.3.: **Filters and nonlinearities of both classes of cells.** **a**, **c** Filters of a derivative-like, tonic subtype of BSN1 (BSN1t) and a leading-suppressive AN3 (STA - red, STC - orange, derivative of STA - gray). The STC filter of the cell in **a** resembled the negative derivative of the STA filter. The STC filter in **b** did not resemble the negative derivative of the STA filter. However, it led the STA filter. **b**, **d** Nonlinearities for the cells in **a** and **b**. Firing rate is color coded (see colorbars right of each nonlinearity). Numbers in the four corners indicate the relative contribution of stimuli falling in the four quadrants of the nonlinearity to the output of the neuron in percent. **e** Coefficient of correlation of the STC filter and the negative derivative of the STA filter (derivative-like: 0.83 ± 0.11, leading-suppressive: 0.31 ± 0.08, $p = 2 \cdot 10^{-11}$, rank sum). **f** Delay between the peaks of the STC and the STA filter (derivative-like: 3.0 ± 0.6 ms, leading-suppressive: 7.8 ± 1.9 ms, $p = 4 \cdot 10^{-6}$, rank sum). **g** Relative contribution of stimuli having positive projection values on the STA filter (weight of upper and lower right quadrants) in driving both classes of cells. Most cells of either class were driven almost exclusively by such stimuli, indicating that the STA was excitatory. **h** Comparison of the relative contribution of stimuli exhibiting positive projection values on the STA filters and positive or negative projection values on the STC filter. While derivative-like cells were driven by positive projection values of the STC filter, leading-suppressive cells fired most strongly in response to negative projection values of the STC filter. Thus, the STC filter was excitatory for derivative-like and suppressive for leading-suppressive cells. **i**, **j** Nonlinearity of the STA filter (i) and STC filter (j) for derivative-like (red) and leading suppressive cells (green, mean $\pm$ s.e.m. over all cells of a each class). **h** Logical operations on the output of the STA and STC filter implemented by the two-dimensional nonlinearity. LEFT: template nonlinearities for a selectivity only for the STA, an `AND` and an `AND-NOT` operation (blue areas: no firing, orange: firing). RIGHT: boxplots showing how well each of the three nonlinearities fits the structure of derivative-like (red) and leading-suppressive cells (green).

3.3.3. Contribution of model components to sparse and decorrelated coding of natural stimuli

The analysis of the encoding properties of local and ascending neurons in the grasshopper revealed two different computational classes: all but one local neuron in the data set could be termed "derivative-like"′ cells, while all ascending neurons were "leading-suppressive"′ cells. The structure of both classes of cells seemed to increase transient firing and thus temporal sparseness. In contrast, while the filters of derivative-like cells were similar to each other, the leading-suppressive cells seemed to exhibit a much larger diversity of STC filters. I expected this latter class of cells to be more effective in increasing population sparseness. To relate the computational properties of both classes of cells to sparse coding I used the models to predict responses to a set of natural songs and quantified the contribution of both filters to temporal and population sparseness (see Fig. 3.4 a for an example song).

Temporal sparseness Temporal sparseness describes the tendency to fire in bursts with long stretches of silence in-between. This results in firing-rate distributions with a large concentration of values at zero and a long tail. While the STA models of both classes fired more strongly during the syllables, the STC models responded almost exclusively at the onset of each syllable (Fig. 3.4 c, d). Accordingly, the STC models exhibited generally higher temporal sparseness than the STA models, with the positive effect being most prominent for the delayed-suppression cells (Fig. 3.4 e, derivative-like: 0.44 ± 0.17 and 0.54 ± 0.19 for STA and STC model, respectively, $p = 9 \cdot 10^{-4}$; leading-suppressive: 0.34 ± 0.18 vs. 0.63 ± 0.23, $p = 1 \cdot 10^{-4}$). Thus, only STC models were able to reproduce the level of temporal sparseness estimated from recordings in chapter 2 (0.52 ± 0.19 and 0.57 ± 0.12 for local and ascending neurons, respectively; see Fig. 2.5 b). Hence, the STC filter and the joint nonlinearity of both model classes substantially contributed to temporal sparseness in the network.

Population sparseness Population sparseness was calculated by constructing four-cell populations of random combinations of models belonging to the same class. Population sparseness is high if different cells in a population do not respond with the same pattern to a sound, e.g. by being selective for different stimulus features. Given the similarity of the STA filters, populations of STA models of both classes exhibited comparable levels of population sparseness (derivative-like STA 0.28 ± 0.06, leading-suppressive STA 0.26 ± 0.06, Fig. 3.4 f). However, while incorporation of the second filter did increase population sparseness only weakly—or even decreased it—for derivative-like cells, leading-suppressive cells greatly profited from the inclusion of the second filter (derivative-like STA 0.28 ± 0.06, STC 0.32 ± 0.06; leading-suppressive STA 0.26 ± 0.06, STC 0.42 ± 0.08, $p = 0$, ranksum). While STA models of both classes and STC models of the derivative-like models exhibited population sparseness comparable to those measured for local neurons (0.35 ± 0.05), only the two-dimensional STC model of leading-suppressive cells approached the high values reported previously for the output of the network (0.47 ± 0.03, see Fig. 2.7 b in chapter 2). As most leading-suppressive cells were ascending neurons, the additional feature described by the STC filter and the associated nonlinearity seemed thus to be necessary for the decorrelation

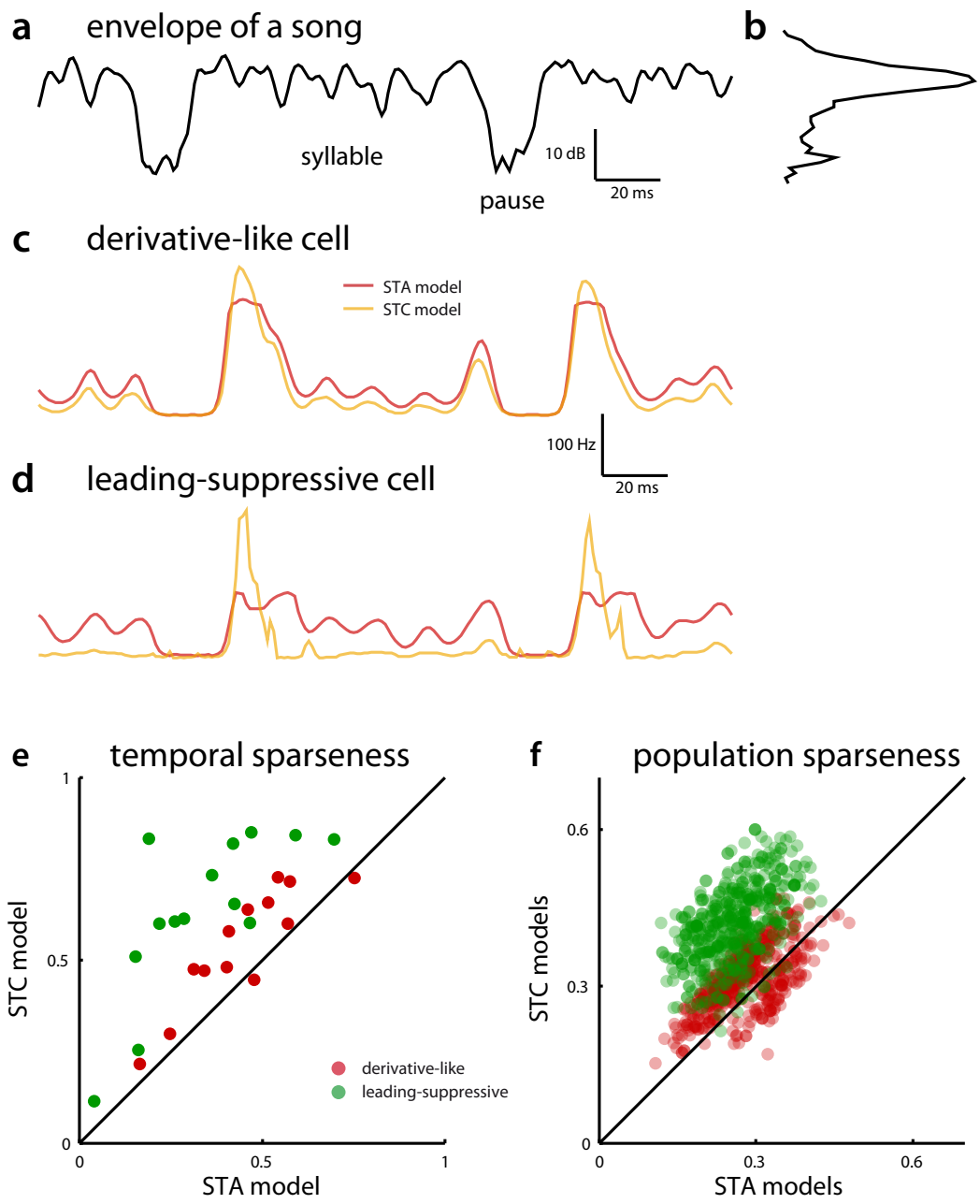

Figure 3.4.: **Contribution of model features to sparse responses. a** Section of the envelope of one of the 8 songs used for simulation. **b** Amplitude distribution of all 8 songs. **c**, **d** Simulated responses of a derivative-like and leading-suppressive model (same as Fig's 3.3a,b and 3.1) to the song in **a**. Red and orange traces correspond to the STA and STC models, respectively. **e** Temporal and **f** population sparseness of the STA and STC models in response to natural songs. Green dots correspond to derivative-like, red dots to leading-suppressive cells. Dots in **f** are semi transparent.

of responses observed in the system.

3.4. Discussion

Employing the framework of linear-nonlinear models, I found two classes of cells in the auditory system of grasshoppers. While the STA filter was similar for both classes, they differed in their second filter: models with a derivative-like STC filter were found at the level of local neurons and models with leading-suppressive STC filter and an `and-not` like nonlinearity were found mostly among ascending neurons.

Simulations have shown that only two-dimensional models reproduce the degree of temporal and population sparseness found in the auditory system of the grasshopper. While both, derivative-like and leading-suppressive cells increased temporal sparseness, only the latter class of cells substantially increased population sparseness.

In the following, I will discuss how the structure of both classes of models increased sparseness and connect the computations described by each model to mechanisms previously shown to increase sparseness.

Temporal sparseness Temporal sparseness increases if transient firing is accentuated and persistent firing is attenuated, leading to responses with short firing events interleaved by long, silent epochs. The two model classes perform this transformation by two different computations (Fig. 3.4): the derivative-like STC filter of most local neurons leads to a differentiation of the stimulus. This amplifies responses to stimulus transients like on- or offsets. The suppressive STC filter of the ascending neurons also accentuates onsets. In addition, it quenches prolonged responses a few milliseconds after the STA filter drives the cell.

Both operations can be subsumed under the phenomenon of spike-frequency adaptation, the decrease of neural firing in response to prolonged stimulation (Benda et al., 2001; Lundstrom et al., 2008; Tripp and Eliasmith, 2010). The relation between spike-frequency adaptation and temporal sparseness has been reported previously (Farkhooi et al., 2009; Houghton, 2009). Adaptation has been described in terms of differentiation (Lundstrom et al., 2008) or as a high-pass filter (Benda et al., 2001), transformations that reduce temporal correlations and thereby increase temporal sparseness (Wang et al., 2003; Tripp and Eliasmith, 2010). The STC filter suggests that derivative-like cells use differentiation for adaptation (see Fig 3.1 e, 3.3). Derivative-like, two-dimensional models as described in this study have been found in many sensory systems (Atencio et al. (2008); Sharpee et al. (2011), see also Brenner et al. (2000); Fairhall et al. (2006); Fox et al. (2010); Slee et al. (2005); Kim et al. (2011)), suggesting that this model structure instantiates—in addition to the contribution to temporal sparseness—beneficial properties, like adaptation to stimulus statistics and robust and efficient encoding of time-varying stimuli (Fairhall et al., 2001; Sharpee et al., 2011).

Adaptation can be implemented by cell-intrinsic mechanisms via adaptation currents (Wang et al., 2003) or in a network via synaptic depression, feedback inhibition (Papadopoulou et al., 2011) and slow, feed-forward inhibition (Assisi et al., 2007; Creutzig et al., 2009).

Incorporating previous knowledge about the grasshopper allows one to speculate on the biophysical basis of the abstract computations described by STC models. The derivative-like STC filter is likely to be implemented by both cell-intrinsic adaptation currents as well as feed-forward inhibition: Two of the 4 cell types belonging to the derivative-like cells—

TN1 and SN1—receive only excitatory inputs from receptors
(Römer and Marquart, 1984). The filter in these cells thus likely arises through cell-intrinsic adaptive potassium or calcium currents. The tonic BSN1 with a derivative-like filter receives, in contrast to the phasic BSN1, only weak inhibitory input. Its STC filter is thus likely to be dominated by cell-intrinsic adaptation currents as well. The other derivative-like cell type—SN3—receives slow inhibitory inputs, which could thus—in addition to intrinsic currents—contribute to the shape of the STC filter (Stumpner, 1988). Note that both inhibitory as well as adaptation currents are all hyperpolarizing currents which move the membrane potential away from threshold and thereby decrease the tendency to fire.

Derivative-like STC filters can also be the result of imprecise locking of spikes to the STA filter Dimitrov et al. (2006). However, this is only likely if such spike-time jitter is greater than the time-scale of the filter (see e.g. Fairhall et al. (2006); Sharpee et al. (2011)). Here, the jitter is smaller than the width of the filter. I calculated spike-time jitter as the standard deviation of the timing of individual firing events across trials as in Desbordes et al. (2008). The filter width was always much larger than the spike time jitter, for derivative-like 10 times and for leading-suppressive cells 6 times larger (jitter: derivative-like 0.6 ± 0.4 ms, leading-suppressive 1.3 ± 0.5 ms; filter-width derivative-like 5.3 ± 0.9 ms, leading-suppressive 6.8 ± 1.7 ms; see e.g. Fig. 3.1 e for an example response and Fig. 3.2).

In the leading-suppressive cells, temporal sparseness increases by shutting-off persistent responses via slow suppression. The two properties of the leading-suppressive models contributing to this transformation are the delay between the STA and the STC filter as well as the `and-not` like nonlinearity and are readily implemented in a network. The STA filter corresponds to excitatory inputs and drives the cell; as the STC filter leads the STA and is suppressive, the cell will only fire strongly if the stimulus preceding the STA is relatively soft. This is equivalent to a slow inhibition suppressing persistent firing by kicking in shortly after the excitation. Such an implementation is highly likely for the phasic BSN1, AN1 and AN3, for which strong, slow inhibitory inputs have been shown in dendritic recordings (Römer and Marquart, 1984; Hildebrandt et al., 2009; Stumpner, 1988; Römer and Seikowski, 1985). For the other leading-suppressive cell type in the data set—AN2—a strong after-hyperpolarization has been shown to underlie adaptation (Hildebrandt et al., 2009); yet, this cell also receives strong contralateral inhibition that can be slower than the excitation (Stumpner, 1988). The STC filter of AN2 is thus likely to be a combination of both cell-intrinsic adaptive currents and inhibitory inputs.

Hence, incorporating prior knowledge about the physiology of the cells in the dataset allows to propose putative biophysical mechanisms implementing the abstract computations described above: temporal sparseness increases in derivative-like cells primarily via cell-intrinsic adaptation, in leading suppressive cells via the interplay between fast excitation and slow inhibition. This reflects an increase of the complexity of involved mechanisms in subsequent stages of the auditory system of the grasshopper: All derivative-like cells were local neurons which pool the responses of receptors and gain their sensitivity to the derivative via a cell-intrinsic mechanism. In contrast, most leading-suppressive cells are ascending neurons; their properties are likely dominated by inhibitory and excitatory synapses from local neurons, and are thus dominated by a network mechanism.

Population sparseness The results show that differentiation of the stimulus and slow inhibition increase temporal sparseness by reducing persistent firing (Fig. 3.4 c–e). This property in itself does not necessarily lead to population sparseness. For population sparseness to be high, cells in a population need to exhibit little tendency to fire together by being selective for different features of a stimulus.

The ability of derivative-like filters to increase population sparseness was relatively small (Fig. 3.4 f). As the STA filters were relatively similar and the STC filter was heavily constrained by the shape of the STA filter—on average 0.83 % of the STC shape of each cell was explained by the STA filter—this second filter varied only little across cells (see Fig's 3.2 a and 3.3 e). Accordingly, the cells comprising the class exhibit a very similar feature selectivity and respond thus uniformly to a stimulus.

In contrast, the STC filter of leading-suppressive cells strongly increased population sparseness up to the values observed in the auditory system of the grasshopper (Fig. 3.4 f). The highly diverse feature selectivity in these cells is established through a large range of delays between the excitatory STA filter and the suppressive STC filter (between 4 and 11 ms, Fig. 3.3 f). As argued above, the model structure of leading suppressive cells is mostly generated by feed-forward inhibition where the inhibition lags the excitation (see also Luo et al. (2010)). The role of excitation and inhibition in shaping temporal filters and in decorrelating responses between cells in a population has been appreciated previously (Schmuker and Schneider, 2007; Wiechert et al., 2010; George et al., 2011). In addition to the filters of leading-suppressive cells being diverse, the `and-not` like joint nonlinearity equips these models with a highly nonlinear operation to select a small set of stimuli for firing (Fig. 3.3 d, h). This narrows the tuning of leading-suppressive cells and reduces the overlap between the responses of different cells of this type. The `and-not` like computation also leads to a "delayed anti-coincidence detection"—the cells fire strongly only if the stimulus at different delays is *not* loud; this can yield a combinatorial and synergistic code (Osborne et al., 2008; Schneidman et al., 2011).

Conclusion This chapter has shown that two-dimensional computations underlie the development of a temporally and population sparse representation of sounds in the auditory system of grasshoppers. Interestingly, ascending the network did affect the basic shape of the STA filter only little. However, the second filter equips neurons with transformations that decorrelate responses in time and in a population. Additionally, the two-dimensional nonlinearity allows cells to specifically select a small subset of the feature space spanned by the STA and the STC filter. Both, the additional filter and the more selective nonlinearity, are factors that contribute to a transformation from a relatively uniform representation in second-order neurons to a diverse population code for temporal features of the song in third-order neurons.

Mechanisms implementing these abstract computations are ubiquitous in nervous systems; the results found in the grasshopper are thus likely to constitute general principles underlying the transformation of neural representations.

4. A model of song evaluation in grasshoppers

In this chapter, I will ask what possible benefits the specific, labeled-line like representation of song in grasshoppers could have for the behavioral evaluation of song. To that end, I will design a read-out circuit that ignores the "when" of temporal features of song but relies only on their "what" as hypothesized in chapter 2. I will show a method to learn behaviorally relevant features from behavioral data and test its ability to predict the behavioral responses to novel stimuli. Furthermore, I will use the structure of the trained read-out circuit do learn about the nature of behaviorally relevant features and their integration.

4.1. Introduction

Finding stimulus features that drive behavior is one of the main goals of neuroscience. This can be easy for simple artificial stimuli when only a single parameter is changed. However, this artificial approach is also unrealistic, as in natural stimuli multiple parameters often covary in a complex manner (Felsen and Dan, 2005; Brette, 2010). Due to the rich correlation structure and high dimensionality of natural stimuli, the relation between stimulus parameters and behavior under natural conditions is often elusive and not easily comprehensible with simple tuning curves.

If a tight temporal relation between a stimulus feature and the behavioral response exists, reverse correlation methods yielding so-called classification images or perceptive fields have been applied with great success (Murray, 2011; Geisler et al., 2009; Neri and Levi, 2006; Neri, 2004; Ahumada and Lovell, 1971; Shub and Richards, 2009). This approach is related to the spike-triggered average, except here, each stimulus is weighted by the behavioral response it evokes and not by the firing rate.

Unfortunately, behavioral responses are often loosely timed with respect to the features controlling them: response timing can be imprecise due to noise, response latency often correlates with stimulus parameters, or a response may be triggered only after the full, non-stationary stimulus was evaluated. All these confounding factors lead to weak or no temporal correlation between a relevant stimulus feature and a behavioral response for a time-varying stimulus, as the feature might have occurred anytime during the stimulus.

Similar problems arise when studying the song recognition system of grasshoppers of the species *Chorthippus biguttulus*: Male grasshoppers produce a calling song. Female grasshoppers decide based on temporal features of the song's envelope whether they should respond or not, initiating a call-response cycle which may eventually lead to copulation (von Helversen, 1972). The male song consists of 20-30 repetitions of a basic subunit—a syllable-pause pair (Fig. 4.1 a i). As females usually wait for the song to end before they respond, there exists no tight temporal relation between relevant stimulus features occurring during the songs and the behavioral response.

The relation between stimulus features and behavior in grasshoppers has been studied by measuring behavioral tuning either for features of the Fourier spectrum of the envelope or for temporal features of the envelope. Both methods constitute pre-processing steps which yield a representation of the song that reduces the temporal dimension of the song by averaging the output of feature detectors over time.

E.g. the power-spectrum of the envelope of the sound represents the average energy each modulation frequency has over the whole song. In addition to providing translation-invariance, it also constitutes a simple algorithm for evaluating long, periodic patterns (Hennig, 2003, 2009; Schul et al., 1999). Indeed, the simple and species-specific structure of the power spectra of the envelopes has inspired such a hypothesis (see Fig. 4.1 a ii). However, careful behavioral experiments have shown that the power spectrum alone cannot explain behavioral selectivity of female grasshoppers (von Helversen and von Helversen, 1998; Schmidt et al., 2008). Rather, information inherent in the phase-spectrum, which dictates temporal features like the shape of transients, is important for perceptual decision

making (see e.g. McCotter et al. (2005)).

While song evaluation based on the power-spectrum alone is thus unlikely, paradoxically, some experiments indicate that the timing of features within the song is not important: as long as the average duration of all pauses during a song has the correct duration, the song will be attractive, irrespective of their number and individual duration (von Helversen and von Helversen, 1998; Creutzig et al., 2010). Thus, the "when" of pauses is not important, only their time-average "what". Taking both results together and phrasing them in terms of the Fourier spectrum of the envelope: while the global phase of features—their timing within the song ("when")—may not be important, their local phase—determining the shape of the temporal feature ("what")—clearly is.

This, together with previous knowledge about the influence of individual temporal features provided by experiments with simple, artificial stimuli has inspired Wittmann et al. (2011) to define a set of 8 temporal features that were thought to be relevant for song recognition. The average feature values over the whole song were then used to train an artificial neural network to predict the response to novel patterns. While this approach yielded high prediction performance and thus indicates that this set of 8 stimulus features is sufficient to describe song recognition, it provided only limited insight into *how* these features are actually extracted by the animal.

Here, I present another approach, inspired by the results on sparse coding in the grasshopper obtained in chapter 2. I have shown that the early auditory system seems to trade information about the timing of features—their "when"—for explicit information about their nature—their "what". I have described the output of the early auditory system of the grasshopper as a bank of feature detectors, which signal the values of different temporal features of the song in their spike count, while the spike-timing is relatively imprecise. Thus, while information about the global phase of features—their timing—might be lost to the nervous system, information about local phase spectra is still encoded in the form of the features different neurons respond to. The song recognition centers in the grasshopper's brain might thus only evaluate the average spike count of particular feature detectors for a song. Thereby, behavior would become invariant to the global timing of features while still being sensitive to the local phase spectrum (see also Creutzig et al. (2010)). Such a mode of song evaluation would be a simple solution to the analysis of long and redundant signals—as exemplified by many insect mating songs—by a nervous system and would account for the seemingly conflicting behavioral evidence presented above.

To test this hypothesis directly, I implemented a classifier that incorporates knowledge about the early auditory system of grasshoppers. The algorithm consists of one or more feature detectors—here implemented as linear filters each being followed by a sigmoid nonlinearity (compare chapter 3)—whose average output over the stimulus is then linearly transformed to the behavioral score of interest. The classifier does not implement any *ad hoc* assumptions about the nature of the features influencing behavior but learns them directly from the behavioral data.

Below, I will describe the classifier and its estimation in more detail. I will then show that it successfully predicts behavioral responses for different data sets obtained from the grasshopper. In contrast to highly non-linear, "black-box" classifiers like artificial neural net-

works, the simple structure implemented here provides insight into what features the classifier is based on and how multiple features are integrated to yield a behavioral response. Furthermore, I will discuss to what extent the learned classifiers can inform hypotheses about neural implementations.

4.2. Methods

4.2.1. Layout of the classifier

The classifier consists of four stages (see also Fig. 4.2):

1. **Extraction of the envelope:** The pattern of amplitude modulations or envelope of a signal was extracted by the root mean square method from the stimuli. It was then transformed to the dB scale and thresholded at 35 dB. This signal formed the input $s(t)$ to the classifier.

2. **Feature detection:** $s(t)$ was linearly filtered $f_j(t) = \int_{-\infty}^{\infty} s(\tau)h_j(t-\tau)d\tau$ and subsequently transformed with a sigmoid nonlinearity $g_j(t) = 1/(1+\exp(-a_j f_j(t) - b_j))$. The combination of the filter h_j and the nonlinearity g_j constituted the feature detector and was similar to linear-nonlinear models of single neurons (compare chapter 3).

3. **Feature integration:** The average output of the feature detector over the stimulus gives the feature value: $v_j = 1/T \int_0^T g_j(t)$. Thereby, the classifier looses all information about the specific timing of the features encoded by the feature detector.

4. **Transformation to behavioral output:** Finally, linear regression on all features yields the classifier's prediction of the behavioral response: $\hat{y} = w_0 + \sum_j w_j v_j$.

4.2.2. Training

The classifier was trained on 4 different data sets for which behavioral scores have been determined experimentally. For details on the stimuli and the procedures to quantify response probabilities see below. The first stage of the model is fixed. The threshold was set to 35 dB based on measurements of primary auditory neurons in grasshoppers (Machens et al., 2001). Pre-tests with thresholds of 30 and 40 dB yielded similar results (data not shown). The envelopes of the full ensemble of stimuli in each data set were normalized to have zero-mean and unit-variance. Parameters in stage 2—the linear filter $h_j(\tau)$ and the parameters for the nonlinearity a_j and b_j—are learned using a genetic algorithm (for implementation details see below). Filters f_j were represented as a weighted sum of up to 16 raised cosine basis functions (Pillow et al., 2008). All 16 components yielded a filter with 64 ms duration. This sped-up training time by reducing stimulus dimensionality and enforced smooth filters. The dependence of classification performance on the duration and number of the filters was determined by learning filters with varying numbers of components or different numbers of filters, respectively. The regression weights w_j in step 4 were determined by standard linear regression.

The genetic algorithm. A genetic algorithm was used to optimize the parameters of the filter and the nonlinearity (step 2). Genetic algorithms are a class of swarm optimization methods which are inspired by biological evolution: A population of solutions is randomly initialized. Then, the fitness of each individual solution is evaluated using a goodness-of-fit

measure between the predicted and the measured behavioral response. Individual solutions are propagated to the next generation based on their fitness. This way, the best solutions have a higher chance to surviving. Variability is introduced to each individual solution through random mutations or recombination of existing solutions. Thereby, good solutions are modified to eventually produce better solutions. This evaluation–selection–mutation cycle is repeated until an optimal solution is found.

In the following I will provide implementation details to enable reproducibility of the results. However, I do not give a comprehensive introduction to the implementation of genetic algorithms. For this see Mitchell (1998).

Training started with a set of 500 random solutions. Model parameters were encoded as 8 bit, binary strings. For a single filter model that has 16 filter coefficients and two parameters for the nonlinearity, each genome is thus represented by a string of $(16 + 2) \cdot 8 = 144$ zeros or ones. The evolutionary algorithm worked on this binary representation for mutation and recombination. For evaluation of the fitness, each “genome" was converted to a decimal representation. Filter coefficients were rescaled, such that they could assume values between -1 and 1. The steepness parameter of the nonlinearity a_j was scaled to range between 0 and 32 (thus constrained to have positive slope), the threshold parameter of the nonlinearity b_j was scaled to range between 0 and 12. Constricting the parameters of the nonlinearity reduced the degeneracy of the fitness landscape: Scaling of the filter can be compensated for by an appropriate rescaling of the nonlinearity, producing two equivalent solutions which reside in disjunct regions of the parameter space. Also, inverting the filter by multiplication with -1 is equivalent to mirroring the nonlinearity along the $y = 0$ axis. Pre-tests have shown that enforcing positive slopes and a limited dynamic range reduced this ambiguity without constraining the power of the trained classifier.

Each individual solution’s ability to reproduce the behavior, their fitness, was quantified by a measure derived from the mean-squared error: $1 - \sum_i (y - \hat{y})^2$, were y is the measured behavioral score and $\hat{y}$ the predicted one and the sum is taken over all stimuli in the training set. The fitness values were scaled such that their mean and standard deviation was 1.0—a procedure called σ scaling. This equalizes selection pressure over the duration of the optimization process. The 10 best solutions out of the 500 in a population were propagated unmodified to the next generation in order not to loose good solutions and to retain a pool of “elite” solutions for recombination. The remaining 490 individual solutions of the next generation were filled up with solutions from the current generation based on their fitness, that is the higher the relative fitness of an individual solution, the more likely it was to be found in the next generation. This selection was implemented via the stochastic universal sampling algorithm, which ensures that solutions are propagated to the next generation according to their relative fitness. Solutions were then modified through mutation and recombination to introduce variability. The mutation rate was 0.003 bit/generation, meaning 3 out of 1000 bits in the binary representation of the parameters were flipped randomly from 0 to 1 or vice versa in each generation. Thus, in a population of 500 single-filter classifiers, each having a genome of 144 bits, 217 bits are flipped in each generation. Recombination was performed by bit-wise cross-over with a probability of 0.8 (also called uniform cross-over); that is, genes between randomly chosen pairs of “parent” solutions are exchanged to

produce two new "child" solutions. The algorithm was terminated after 500 generations or if the fitness of the best solution in the population was within 0.001 % of the population's mean fitness.

Quantification of classifier performance: Training was performed with 100 runs of a 10-fold, stratified, hold-out cross validation. That is, the classifier was trained on 9/10th of the stimuli and then tested on the 1/10th of stimuli not used for training. It was ensured that the training and test data covered roughly the same range of behavioral response values ("stratified" sampling). Performance was quantified using two measures: The first is the mean squared error (mse) $\left\langle (y - \hat{y})^2 \right\rangle$, where angled brackets denote average over all stimuli in the test set. The second measure is Pearson's coefficient of correlation r^2. While the mse gives a measure of the magnitude of the deviations of predicted responses from the measured ones, r^2 is the variance in the measured response scores explained by the predictions. As the latter measure depends on the variance of the data, it can yield bad scores in the case of low variance in the data despite the magnitude of errors being low. If not stated otherwise, all performance values given are the average test performance over all cross-validations.

4.2.3. Data for training and testing

To test the ability of the proposed approach to predict the attractiveness of stimuli for female grasshoppers, the classifier was trained on four different sets. For these four data sets, behavioral responses have been measured in playback experiments as described in Schmidt et al. (2008). Shortly, a female grasshopper of the species *Chorthippus biguttulus* was placed in a sound-proof chamber. Playback of signals, and the recording and detection of female response songs was controlled by a computer. All signals in a set were presented 18 times, in randomized order. General motivation to engage in song recognition was tested by presenting an attractive, block-like pattern (block-like, 80 ms pulse, 12 ms pause, 70 dB plateau). Stimulus presentation was halted until the female responded to this positive control. The probability of the female to respond to the presented stimulus was taken as the behavioral response value or attractiveness of a particular stimulus. This measure was then normalized by the response probability to the positive control stimulus interleaved in the normal testing procedure.

The four different stimulus sets on which the classifier was trained differed in the pattern of amplitude modulations; the carrier spectrum was identical across all data sets and consisted of band-limited white noise with power between 5 and 40 kHz. The stimulus sets were originally designed to investigate different aspects of song recognition in grasshoppers: The first two sets consisted of artificial, block-like stimuli constituting a behaviorally very effective abstraction of the syllable-pause structure of natural songs (Fig. 4.1 b). The first set (data set A) tested the influence of plateau intensity and an accentuated syllable onset on the behavioral tuning for pause duration (Fig. 4.1 b i, ii, iii, kindly provided by Jana Sträter). Pause duration is a species-specific temporal parameter of male calling songs and female grasshoppers usually exhibit a band-pass tuning centered around the conspecific pause duration. The second set (data set B) also probed the influence of temporal parame-

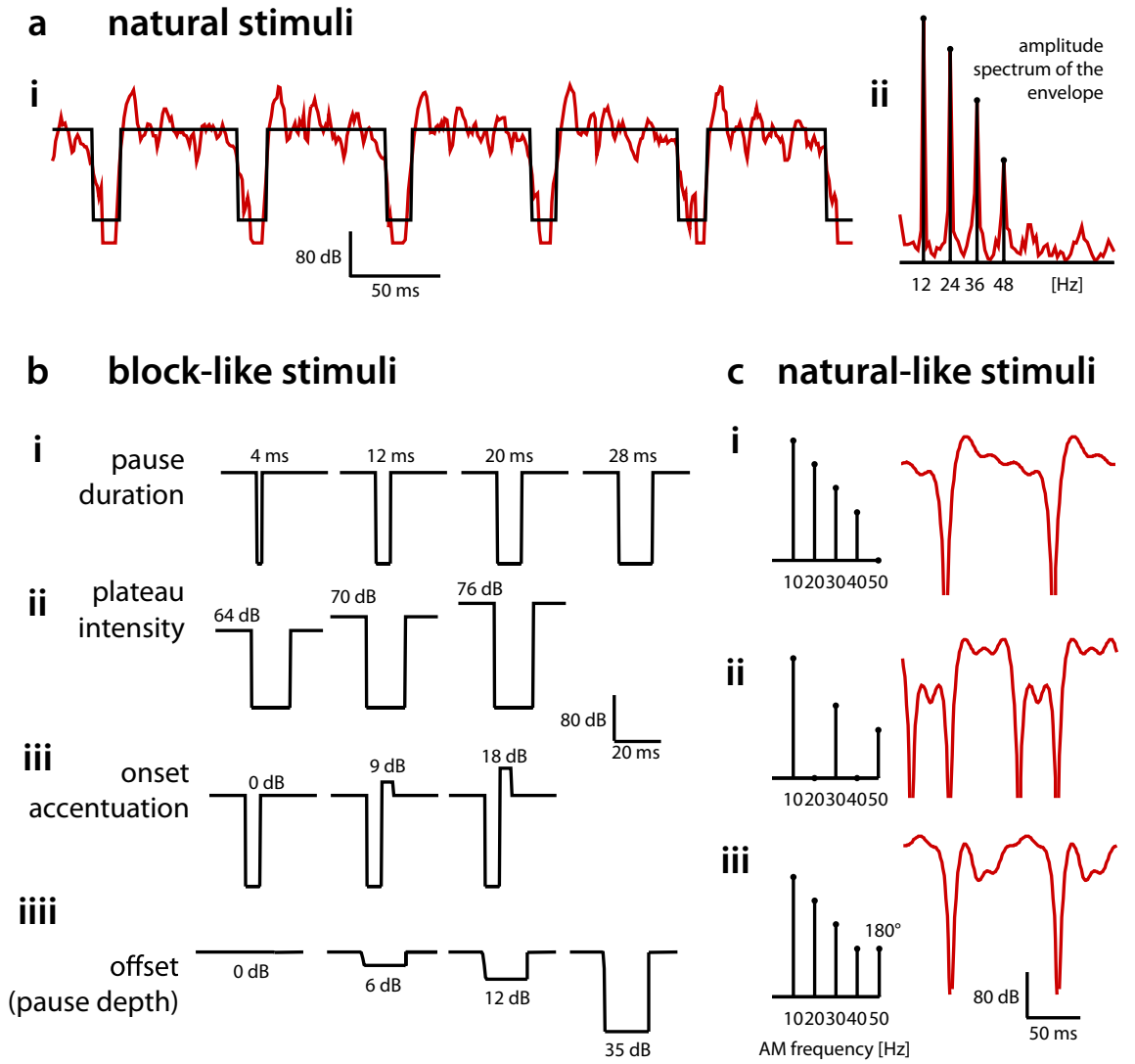

Figure 4.1.: **The four stimulus sets used for training and testing the classifier. a i** Short section of the envelope of a natural song of a male grasshopper (red). The song consists of many repetitions of a simple subunit: a loud syllable followed by a soft pause. Often, the onset of the syllable is accentuated. The natural song can reduced to a block-like pattern, while retaining its attractiveness (black). **a ii** The amplitude spectrum of the full song from the left exhibits a simple harmonic structure with a fundamental at 12 Hz and three harmonics with decreasing power. **b** A great variety of parameterizable temporal patterns can be defined by starting from a simple block-like stimulus and changing the plateau amplitude, the pause duration, the pause depth (offset) or the height of the onset accentuation. Onset and offset are defined relative to plateau intensity. **c** By modifying the amplitude and phase of a 10-Hz-fundamental and its 4 harmonics, a multitude of complexly, natural-like patterns can be created.

ters on the tuning for pause duration. However, here, the influence of pause depth (or offset) and onset accentuation was tested (Fig. 4.1b i, iii, iiii, kindly provided by Olaf Kutzki). The third set (data set C) consists of stimuli being built from a Fourier domain representation of the envelope of conspecific natural songs (Fig. 4.1a ii). They were constructed from linear combinations of sines with a frequency of 10, 20, 30, 40 and 50 Hz, different amplitudes and phases. This set was originally presented to obtain insights into whether the evaluation of song can be explained by the power spectrum of a signal alone (Schmidt et al. (2008), Fig. 4.1c, kindly provided by Arne Schmidt and Matthias Hennig). The stimuli exhibit a rich temporal structure where many features of the song change in an often correlated manner. The fourth set (data set D) contains 40 different natural songs recorded from different male grasshoppers collected at several spots in Germany (Fig. 4.1a, kindly provided by Nicole Stange).

4.3. Results

The classifier proposed here consists of four processing stages which can be identified with elements of the auditory system of grasshoppers (Fig. 4.2): In the input stage the envelope of the song is extracted, thresholded and transformed on a dB scale. This first stage is implemented in auditory receptors, which encode a faithful representation of the signal's envelope, linear on a dB scale (Machens et al., 2001; Gollisch et al., 2002). Then, temporal features of the song are detected in parallel by up to three filter-nonlinearity cascades. Second- and third-order neurons in the early auditory system of the grasshopper are well approximated by such linear-nonlinear models (see chapter 3). Feature values are obtained by averaging the output of each detector over the full stimulus. This removes all information about the timing of the features within the stimulus. The feature values are linearly combined to obtain the response probability for that stimulus. These last steps are likely to be implemented in decision centers in the brain.

4.3.1. Performance for individual data sets

I trained the classifier using a genetic algorithm on four different data sets (see Methods for details). To assess the influence of the duration and number of features on classification performance I trained a classifier with 1 to 3 feature detectors (filters and nonlinearities) and filter durations between 6 and 64 ms. For all data sets, a single filter predicted behavioral responses well (Fig. 4.3 a–d). While addition of a second filter improved performance by $\approx$ 20 % for the first two data sets (A, B), it had no effect on prediction for the natural-like and natural data sets (C, D). A third filter never increased performance substantially. Filters exhibiting a temporal support of $\approx$24–48 ms (6–14 components) were sufficient to maximize model performance for one- and two-filter classifiers.

In general, the simple classifier performed well for most data sets (Fig. 4.4). The classifier almost perfectly predicted behavioral responses for the block-like stimuli of data sets A and B (Fig. 4.4 e, f). Furthermore, the prediction for these two data sets remained mostly within the margins of error for the behavioral response score (Fig. 4.4 a,b). The mean squared-error was < 0.004, the coefficient of correlation 0.96 and 0.90, respectively.

For the natural-like stimuli of data set C, performance was in general good (mse 0.025, $r^2 = 0.65$, Fig. 4.4 g). However, the classifier exhibited outliers which exceeded the margin of error for the response (Fig. 4.4 c).

Performance for the set of natural stimuli of data set D appears to have been relatively poor with an r^2 of 0.29 (Fig. 4.4 h). This is mostly because this data set spanned only a small range of response values between 0.4 and 0.85. As r^2 is normalized by the variance of the data, it yields low values for data with low variability even if the absolute error is small. Indeed, the mean-squared error was 0.009 and therewith comparable to those of other data sets. Furthermore, predictions rarely exceeded the margin of error for the response score, mainly because of the s.e.m. being larger than in the other data sets (Fig. 4.4 d). All in all, performance for natural stimuli is comparable to that of the data sets (Fig. 4.4 i).

In the following I will further discuss the structure of the classifier obtained for block-like (data sets A and B) and natural-like stimuli (data set C). The natural songs will not be

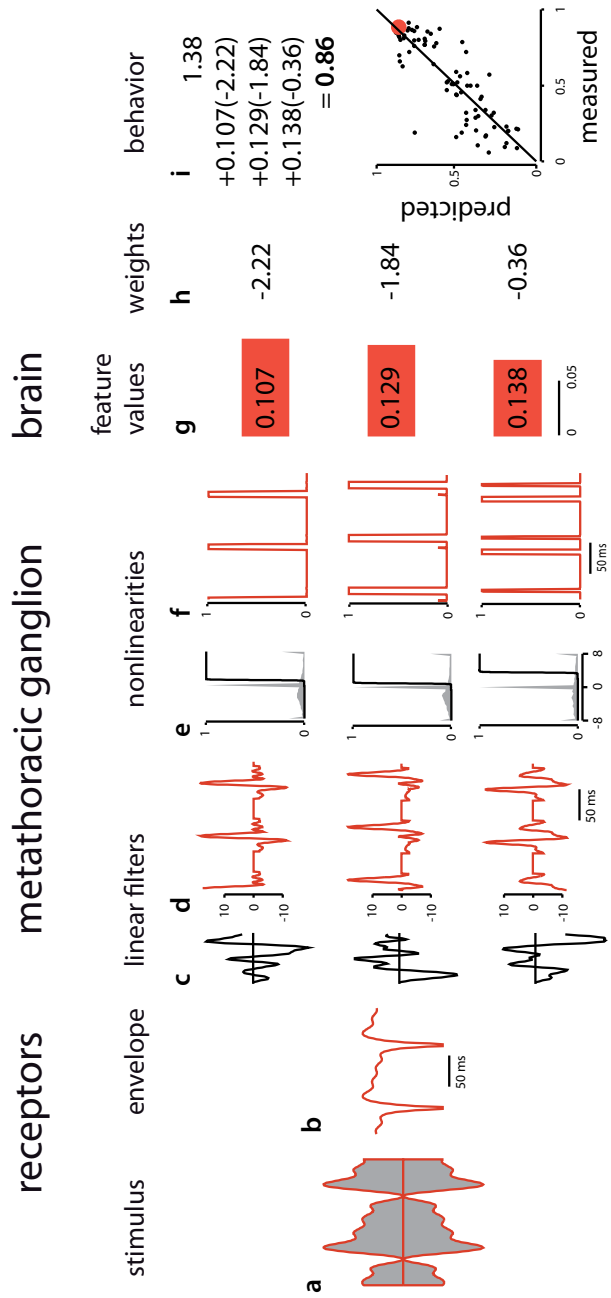

Figure 4.2.: **Structure of the classifier.** **a** The stimulus is a broadband carrier whose amplitude is modulated. **b** Receptors extract the stimulus' envelope and represent it on a dB scale after thresholding at 35 dB. **c** In the metathoracic ganglion, a set of linear filters filter the envelope. **e** A subsequent nonlinearity transforms the filter output (**d**). **g** Feature values are obtained by computing the average output of the nonlinearity (**f**) over the stimulus. **i** Finally, the response probability is computed by linearly weighting (**h**) the feature values. The linear filters and the nonlinearities (**c**, **e**) were trained using a genetic algorithm. The weights—one for each feature plus y-intercept (first term) (**h**)—were obtained by linear regression.

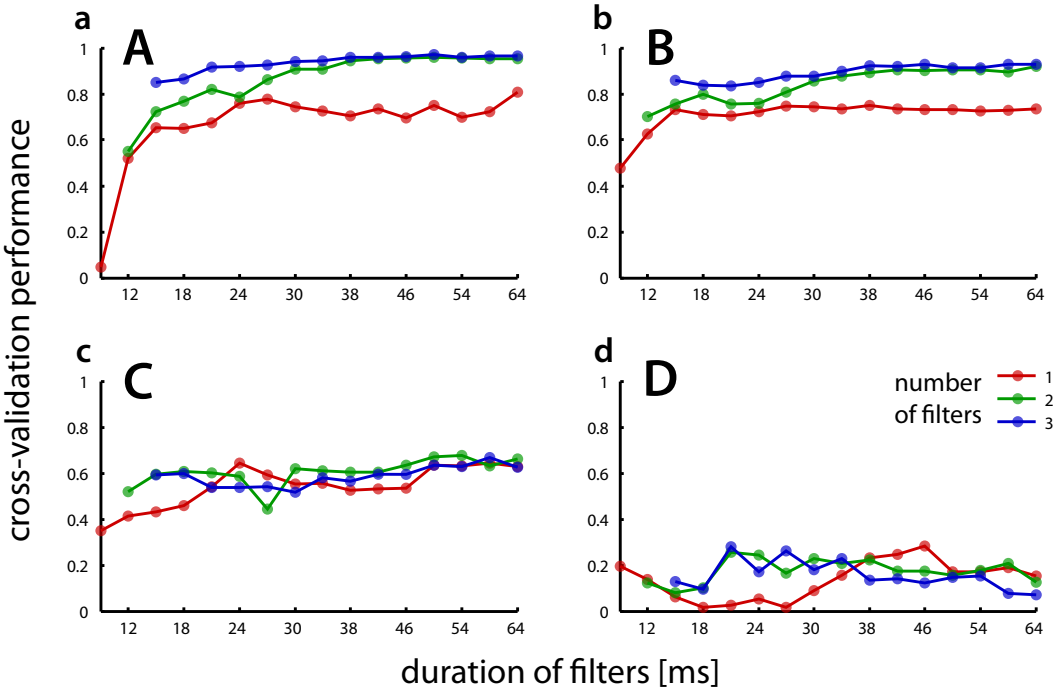

Figure 4.3.: **Dependence of performance on the number of filters and their length. a–d** Cross-validation performance of the classifier for the four data sets (A-D).

discussed further (data set D) due to limited range of response values.

4.3.2. Song-recognition in a block world

One of the main advantages of this approach is that the classifier has a simple structure that is readily comprehensible. It can thus not only serve as a black box to predict responses to novel stimuli but also yield insights into how the envelope of a song determines its attractiveness. As the performance for block-like stimuli (data sets A and B) was very good and both stimulus sets were relatively similar, I trained a *single* classifier on both data sets. This was done without cross validation in order to avoid having to built a consensus classifier out of the set of classifiers obtained for each round of cross-validation. This approach yielded a classifier with a performance and structure not fundamentally different from that obtained by cross-validation on the individual data sets, indicating that over-fitting was no problem in this case ($r^2 = 0.92$ not cross-validated, Fig. 4.7 f, cross-validated $r^2 = 0.96$ and $r^2 = 0.90$, Fig. 4.4 e, f).

I will first show that the classifier successfully predicts many aspects of the behavioral tuning found in the experimental data and will then examine the structure of the classifier to understand how it does so.

Influence of offset and onset accentuation on pause tuning Pause duration, onset accentuation and offset interacted in a complex and nonlinear manner in the behavioral responses (Fig. 4.5 a, c, e): There existed a tradeoff between onset accentuation and offset as the sum of onset and offset needed to exceed 18 dB in order to elicit strong responses > 0.5 (Fig. 4.5 a). However, too strong offsets of 99 dB reduced attractiveness if the pause was long (Fig. 4.5 c). Furthermore, tuning for pause duration was band-pass with an optimal pause at 12 ms but a width and peak value which often depended on the offset: while the width of the tuning for pause was maximal at 12 dB offset, the peak assumed its greatest values and sharpest form at 99 dB (Fig. 4.5 e).

The model reproduced these interactions well (Fig. 4.5 b,d,f): The trade-off between onset accentuation and offset (Fig. 4.5 b) as well as the reduced attractiveness of deep offsets for long pauses (compare Fig's 4.5 b and d) were also seen in the predicted responses. Also, the optimal pause length of 12 ms and the increase of the sharpness and peak of the pause tuning with offset were captured (Fig. 4.5 f).

Influence of intensity and onset accentuation on pause tuning In contrast to the strongly nonlinear impact of onset accentuation and offset on pause tuning, the impact of intensity was relatively simple: The main effect seemed to be scaling and translation of the curve while its basic shape was only little affected (Fig. 4.6 a, c). This was well captured by the model (Fig. 4.6 b, d). Note also that the asymmetric shape of the tuning for pause duration—with a steep rising and a shallower falling slope—was reproduced. Moreover, the model also captured the fact that for strongly accentuated onsets no pauses were necessary to elicit a behavioral response (Fig. 4.6 c,d, blue line for 12 dB onset accentuation and 0 ms pause

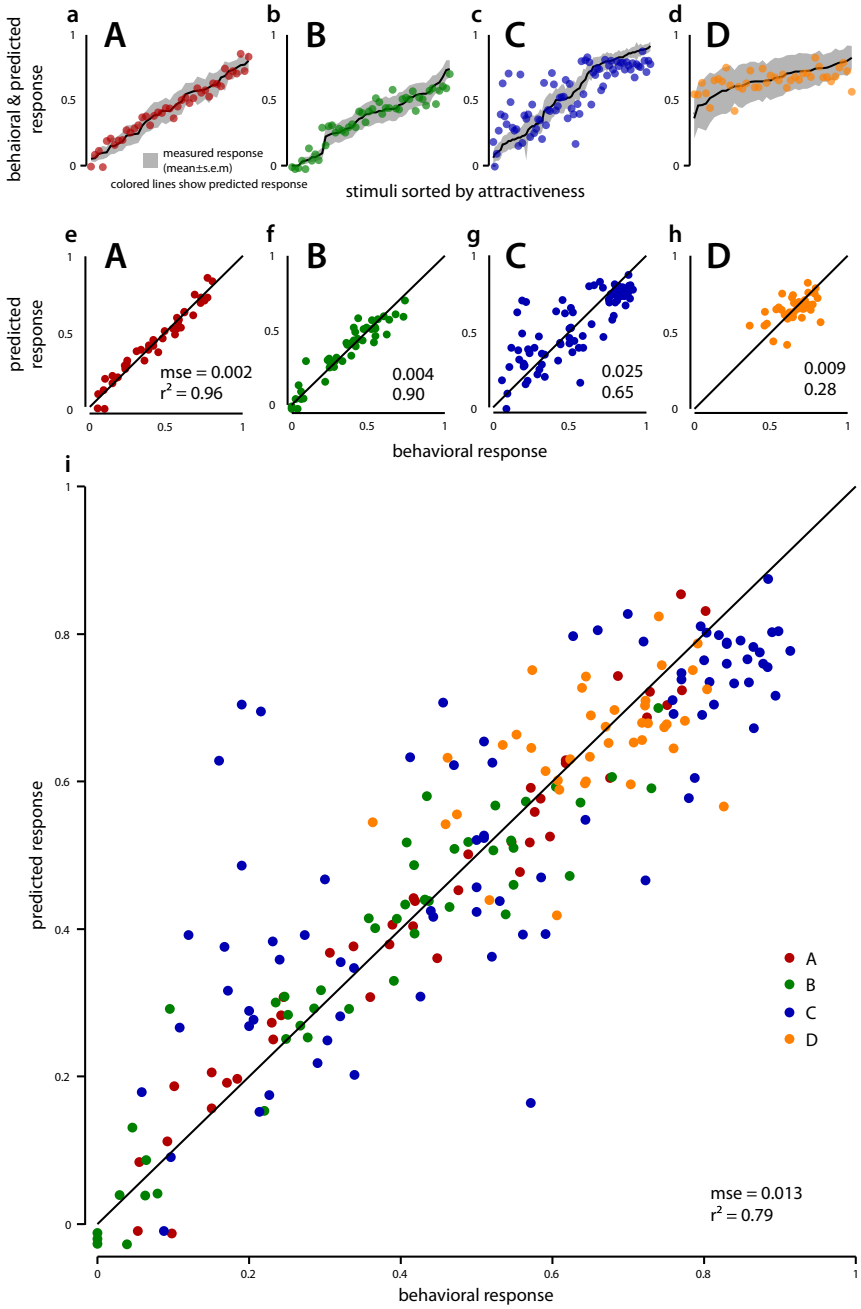

Figure 4.4.: **Classifier predicts behavioral response well for all stimulus sets. a–d** Stimuli *sorted by their attractiveness* are plotted against model prediction. Black line and gray area indicate mean and 5% confidence intervals of the measured attractiveness. Colored points show the predicted response for the respective stimulus. Data sets are indicated as big bold letters. Note that the classifier occasionally predicts negative responses as its output is not bounded between 0 and 1. **e–h** Predicted response plotted against the measured response. Each dot corresponds to one stimulus. Number in the lower right corner show mean squared-error (mse) and correlation coefficient. **i** Measured vs. predicted response for the different classifiers plotted in a single plot.

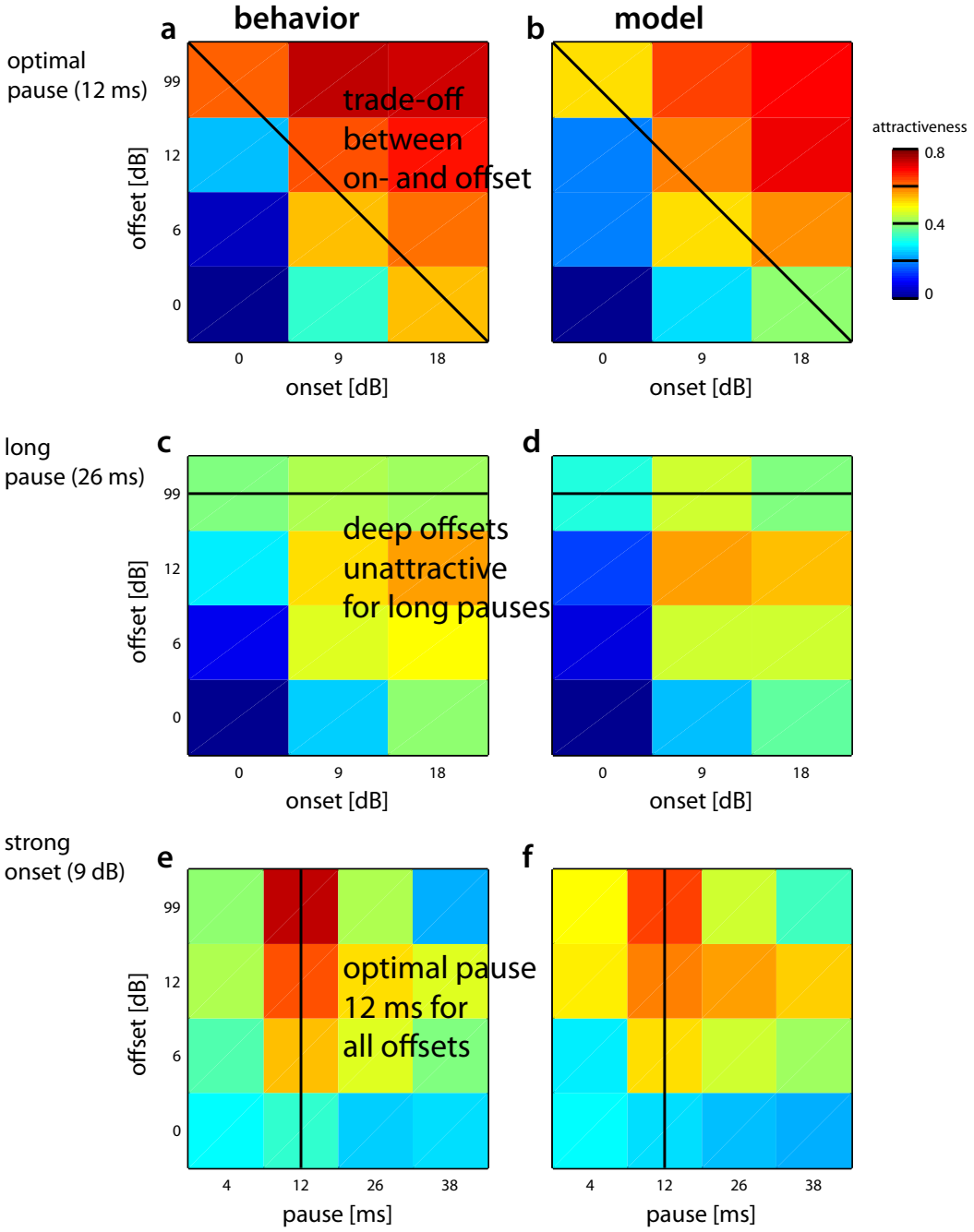

Figure 4.5.: **Dependence of measured and predicted signal attractiveness on pause duration, onset accentuation and offset.** Attractiveness is color coded (see color bar) with cold colors corresponding to low and hot colors corresponding to high values. **a**, **b** Dependence of measured and predicted attractiveness on offset and onset accentuation of a block-stimulus with an *optimal pause of 12 ms*. A trade-off between onset accentuation and offset was clearly visible. **c**, **d** Dependence of measured and predicted attractiveness on offset and onset accentuation of a block-stimulus with a *long pause of 26 ms*. Here, deep offsets became unattractive for all pause duration. **e**, **f** Dependence of measured and predicted attractiveness offset and pause duration of a block-stimulus with an onset accentuation of 9 dB. The shape of tuning for pause duration depended on offset.

duration, see also Balakrishnan et al. (2001)). However, the classifier exhibited a "floor-effect" not seen in the data for short pauses at low intensities, where predicted values did not seem to be able to fall below 0.1 (Fig. 4.6 b,d, red curve).

The fact that the shape of the signal surrounding the pauses strongly influenced the tuning for pause duration and thus that the "perception" of pause duration displayed little robustness, casts doubt on the validity of this parameter for species separation. Either, the pause durations of different species are separated well enough to account for the additional variability in the preference functions introduced by offset and onset accentuation (compare Amézquita et al. (2011)). Or, these latter parameters might as well be differentially expressed in different species. In this sense, the "interactive" relation between pause duration, onset accentuation and offset might constitute a multi-dimensional compound feature and actually support species separation. A detailed analysis of the songs of sympatric species regarding pause duration, onset accentuation and offset might shed light onto that issue. Interestingly, the shape of the tuning for pause duration—though not its scale—was relatively unaffected by plateau intensity (see also Balakrishnan et al. (2001)). As intensity is heavily dependent on the distance between sender and receiver, it is a parameter least indicative about species identity. Having an optimal pause duration that is invariant to intensity is thus advantageous. This invariance in behavior exists despite early processing stages exhibiting relatively little intensity invariance (Clemens et al., 2010).

How do these trade-offs come about?

Given the simplicity of the classifier, it is remarkable that it successfully reproduced the strongly nonlinear and "interactive" tuning observed in behavior. In contrast to many other classification methods (see e.g. Wittmann et al. (2011)), it is relatively easy to dissect the classifier and learn *how* it works. This might provide insight into how song recognition functions on an algorithmic level in the animal.

Recall that the optimal classifier for the block-like stimuli consisted of two linear filters each one with a sigmoidal nonlinearity. The average output of these two linear-nonlinear cascades yields two feature values which are linearly combined to predict the behavioral response to that stimulus. I will first describe the structure of the classifier and then explain how this structure reproduced the behavior.

Unexpectedly, the two feature detectors exhibited highly similar filters (Fig. 4.7 a, $r^2 = 0.84$ between the filters). Both filters consisted of a positive lobe being followed by a negative lobe and responded thus best to offsets of a stimulus such as the end of a syllable or the end of an accentuated onset; responses to the onset of a stimulus would be negative. The negative lobe had a half-width of 6 ms for both filters. This corresponds well to the pause duration at the increasing slope of a general tuning-curve for pause (Fig. 4.6). The positive lobe of both filters lasted ≈ 8 ms.

Although the shape of the filters of the feature detectors was highly similar, the nonlinearities associated with both differed and thus led to different "features" having been detected. Both nonlinearities showed comparable thresholds, responding only to positive outputs of the filters. The first filter exhibited a shallow slope leading to a relatively large dynamic

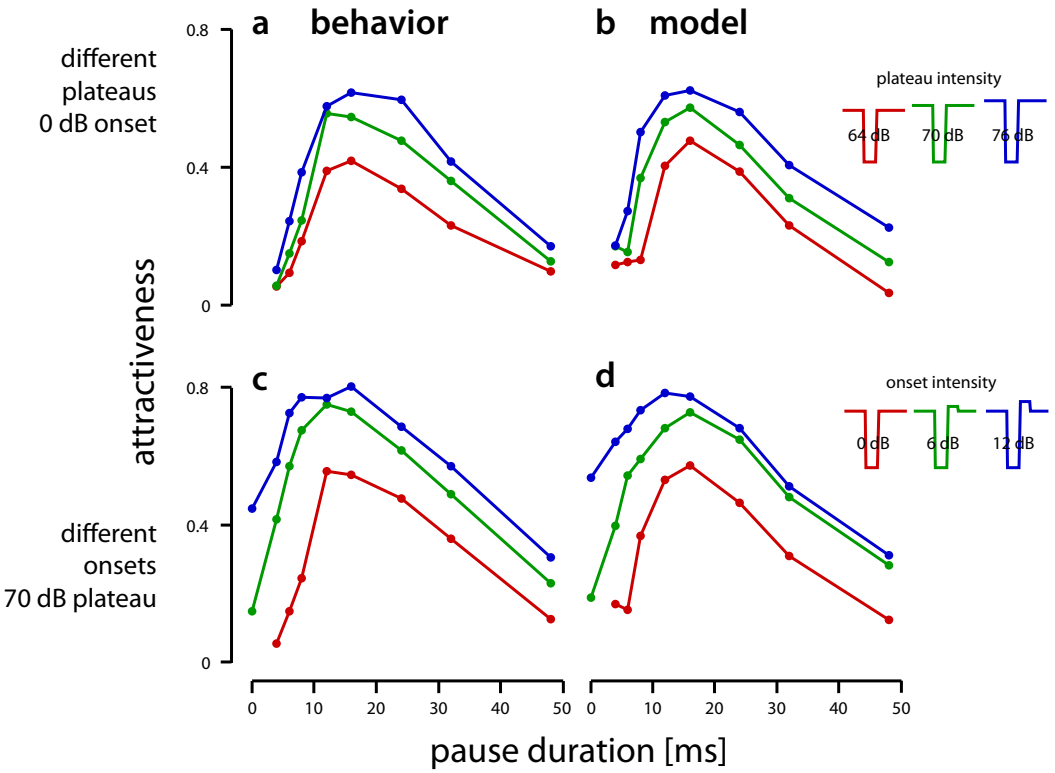

Figure 4.6.: **Influence of intensity and onset on measured and predicted on tuning for pause duration. a**, **b** Pause tuning for block stimuli without an onset and with plateau intensities of 64, 70 and 76 dB (see schematic stimuli to the right). **c**, **d** Pause tuning for block stimuli with onsets of 0, 6 and 12 dB and plateau intensities of 64 dB.

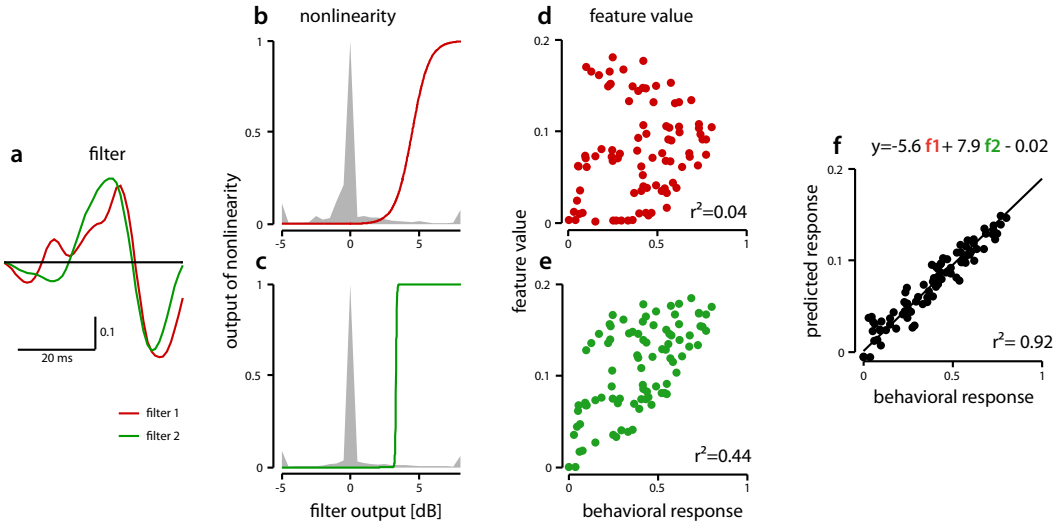

Figure 4.7.: **Two-filter model for block-like stimuli. a** Filters for both feature detectors (red and green respectively). Both filter prefer a stimulus that is first loud and then soft, which correspond to an offset in the signal. **b**, **c** Nonlinearities, transforming the outputs of the filtered stimulus. Grey areas show distribution of output values of the input. **d**, **e** Behavioral response plotted against feature values, which were obtained by averaging the output of the nonlinearity. One dot is one stimulus. **f** Behavioral response plotted against predicted response obtained by linear combination of both feature values (regression formula shown in plot, f1 – feature value of filter 1, f2 – feature value of filter 2).

range (Fig. 4.7 b). In contrast, the second filter exhibited a very steep slope, yielding almost binary responses switching between 0 and 1 depending on whether the filter output exceeded the threshold or not (Fig. 4.7 c). Interestingly, the output of the individual feature detectors correlated only poorly with the behavioral response (Fig. 4.7 d,e, r^2 of 0.04 and 0.44). Only the linear combination of these features yielded a very good performance ($r^2 = 0.92$, Fig. 4.7 f), indicating that the performance of the model was an emergent property of linearly combining two relatively unattractive features. This has implications for searching for the neural basis of behavioral tuning: the individual building blocks (feature detectors, i.e. neurons) need not necessarily be correlated with behavior in order to determine it. The first feature had a negative regression factor of -5.6; high feature values thus suppressed signal attractiveness. The second feature exhibited a positive regression factor of 7.9; high feature values thus increased signal attractiveness.

How was the band-pass filter for pause duration realized? To gain a first intuition about the contribution of both features to signal attractiveness, I determined their tuning for pause duration by looking at their output to stimuli with a 80 ms syllable and pause durations ranging between 0 and 50 ms (Fig. 4.8). This showed that the suppressive first filter was a strict high-pass filter for pause duration while the excitatory second filter was a high-pass filter with a weak band-pass characteristic for long pauses.

Their linear combination yielded a sharply-tuned filter for pause duration with the steep

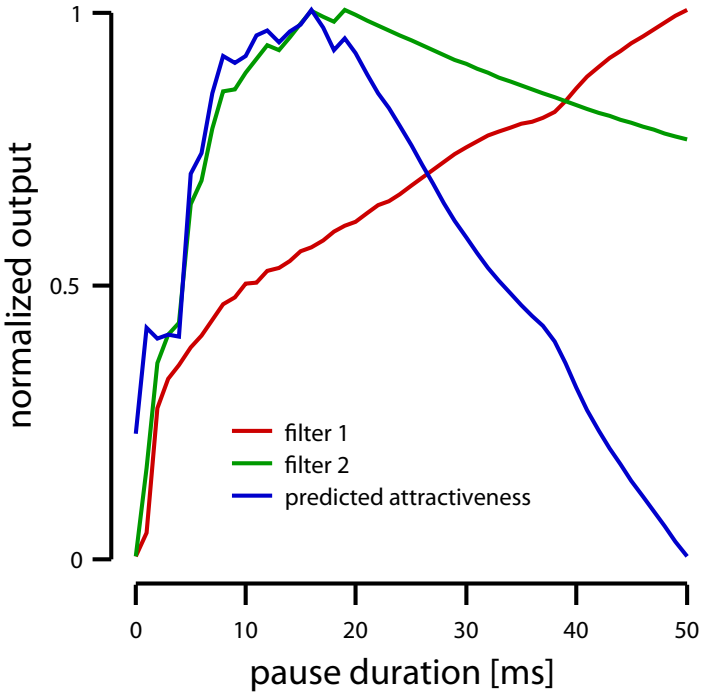

Figure 4.8.: **Tuning for pause duration of both feature detectors and the classifier prediction.** Tuning curves were normalized to range between 0 and 1 to facilitate comparison. The first filter is a high-pass filter for pause duration with a shallow slope while the second filter is a steeply rising, weak band-pass filter. Linear combination of the filters with weights given in Fig. 4.7 f (first filter suppressive, second filter excitatory) yields a sharp band-pass tuning for pause duration.

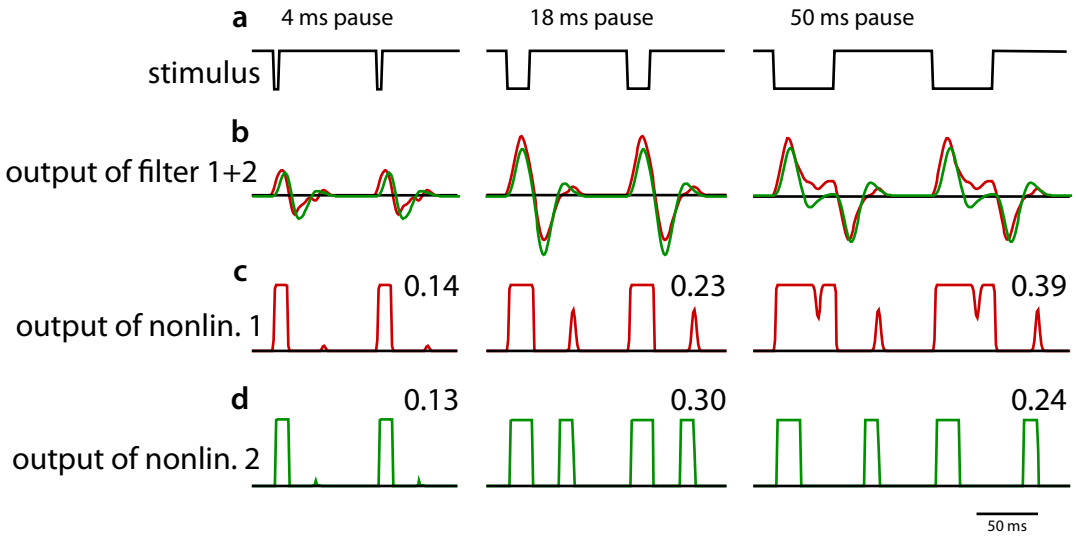

Figure 4.9.: **Pause tuning a** Block stimuli with 80 ms syllable and 4, 18 or 50 ms pauses. **b** Stimuli in **a** filtered with filter 1 and filter 2 (red and green, respectively, filters as in Fig. 4.7 a). The output of the second filter is similar–though not identical—and not shown here. **c** Output of the first feature detector. It responds throughout the pause. Hence, its integral output increases with pause duration yielding a high-pass filter for pause duration (Fig. 4.8, red). Note the short response events during the syllable plateau due to the filter being weakly negative initially and thus responding weakly to an onset. Numbers above the plots indicate feature values corresponding to the time-average of the output of the feature detector. **d** Output of the second feature detector (Fig. 4.7a red, c). For long pauses, the magnitude and duration of output events does not increase. However, as the event frequency decreases for longer pauses, the feature values decreases, yielding the weak low-pass tuning for pause duration (Fig. 4.8, green).

rising and shallower falling flank as seen in the experimental data (compare Fig. 4.6). This setup of filters is akin to those observed in crickets (Schildberger, 1984) and has been proposed for grasshoppers previously (Creutzig et al., 2010). Furthermore, evidence from electrophysiology suggests that such filters might be realized in single, ascending neurons in the early auditory system of the grasshopper (Creutzig et al. (2009), Olaf Kutzki, PhD thesis). Interestingly, there exist individuals which accept very long pauses (von Helversen and von Helversen (1997); Bernhard Ronacher, unpublished results). Their tuning curve resembles that of filter 2 (Fig. 4.8, green), which suggests that in these animals filter 1 probably has little weight as in those rejecting long pauses. The two-filter model of pause tuning has thus the potential to explain inter-individual variability in pause tuning by assuming different weights for the filters.

The rising flank of the tuning curve for pause duration of both filters arose out of the time scale of the filters: Pauses shorter than 6 ms produced weak outputs because of the filter's negative lobe matching the short pause only partly.

The high-pass tuning of the first filter emanated out of the shallow nonlinearity which produced outputs that were sensitive to the duration of the pause and thereby increased with increasing pause duration (Fig. 4.9 c). For short pauses, this was mostly due to the better match of the negative lobe of the filter with the stimulus. For longer pauses, this

was due to the feature detector producing positive output during the pause, yielding larger feature values for longer pauses. In contrast, the second filter exhibited a very sharp, almost step-like nonlinearity with a slightly higher threshold preventing it from responding during a pause (Fig. 4.9 d). As long as the pause was long enough, the output of the filter became invariant to pause duration itself and signaled only the occurrence of pauses. The integral output of this feature detector was thus proportional to the number of pauses in the signal for pause durations > 4 ms. Increased duration of pauses led to a decreased rate of occurrence of pauses. This yielded decreased feature values for long pauses.

Why was there a trade-off between offset and onset? A strong onset accentuation can compensate for a shallow offset/pause, as also onsets exhibited a falling flank at their end—"what goes up must come down". They thus introduced an additional offset and therefore drove the second, excitatory filter. The system was thus sensitive to the overall amount of offsets in the signal. There is evidence that offsets in the middle of the syllable plateau introduced by steps in amplitude render a signal unattractive (von Helversen and von Helversen (1987); Stefanie Krämer, unpublished data). Such steps in the middle of the syllable possibly also drive the first, suppressive filter and thereby reduce attractiveness, similar to the effect of combining very deep offsets and strongly accentuated onsets (see below, Fig. 4.10 b). However, additional experiments and analyses are necessary in order to determine whether this is really the case.

Why did intensity or offset scale attractiveness but did not change the shape or width of tuning? In contrast to on- and offset, intensity had primarily a scaling effect on pause tuning (Fig. 4.6). Looking at the dependence of both feature values on intensity shows why this was so: While the excitatory filter increased with intensity, the suppressive feature decreased yielding a net increase of attractiveness (Fig. 4.10 a, c). As both features scaled linearly with intensity, the resulting tuning curves which were the results of a linear combination of the two features could only exhibit linear changes—that is, scaling and translation (Fig. 4.6). Intensity invariance of pause tuning was thus a result of equally tuned "excitatory" and suppressive features.

Why did very large offsets render longer pauses unattractive? Why did tuning width change with offset? Deep pauses increased the output of both filters as they increased the offset strength (see Fig. 4.5 b). However, the output of the first suppressive filter grew faster than that of the second excitatory filter, eventually surpassing it (Fig. 4.10 b). This reduced attractiveness of long pauses at deep offsets and strong onsets and led thereby to a narrower tuning for pause duration (Fig. 4.10 d). Note, that such an effect could not be achieved with a single filter only, as each individual filter responded monotonously to offset; hence, a single-filter classifier would lack the decrease of attractiveness with deep offsets for long pauses (simulation data not shown).

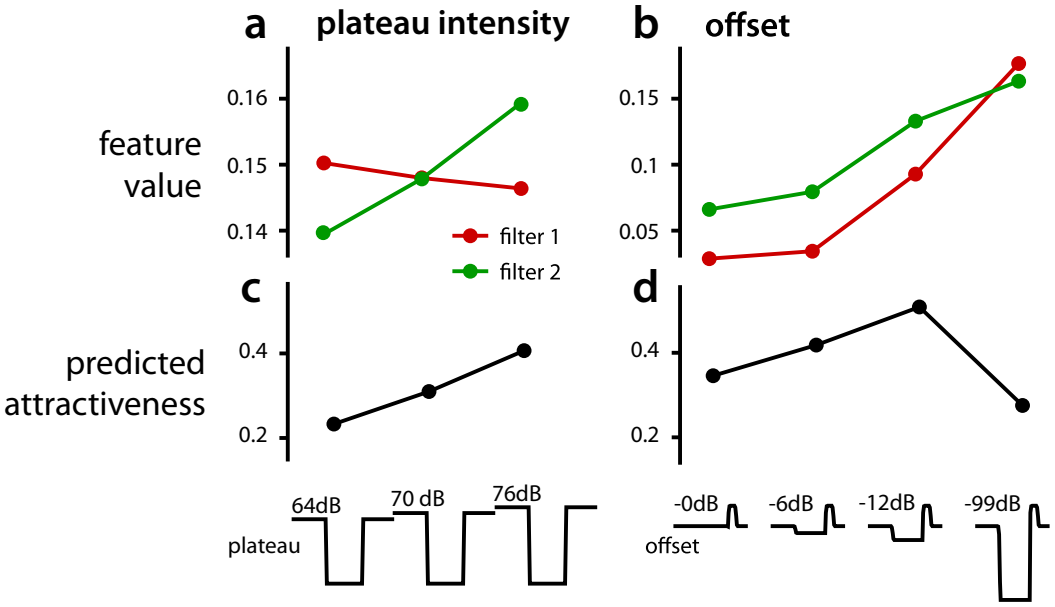

Figure 4.10.: **Effect of plateau intensity and offset strength on behavioral attractiveness of block-like stimuli. a**, **b** Feature values for the two filters (red and green respectively) to stimuli with different plateau intensities (a) or offsets (b). Stimuli are shown below the plots. Patterns in (a) exhibit pause duration of 32 ms and no onset while patterns in (b) exhibit 32 ms pause duration and 9 dB onset. **c**, **d** Predicted attractiveness values, resulting from a linear combinations of both feature values for each stimulus.

4.3.3. Tuning in a natural(-like) world

Above, we have seen that the approach yields valuable insights into how signal attractiveness depends on temporal features for block-like stimuli. The classifier also performed well for natural-like stimuli (Fig. 4.4 c, g). As this data set has been examined in two other studies (Schmidt et al., 2008; Wittmann et al., 2011), I will compare my approach to those taken previously. Furthermore, I will provide an intuition for *why* the classifier proposed here was successful and why it failed for some stimuli.

In contrast to the block-like stimuli discussed above, the stimuli comprising this set are not easily described by a pre-defined set of temporal features. Their envelope was specified by 10 spectral parameters: the amplitude and phase at 10, 20, 30, 40 and 50 Hz. The stimuli exhibited a basic periodicity of 100 ms with a complex fine structure (Fig. 4.1 c).

The most obvious approach would be to relate the spectral parameters to behavior. However, Schmidt et al. (2008) have shown that the relationship between the spectral parameters and behavior is nonlinear. Indeed, multi-linear regression yielded a low performance of $r^2 = 0.47$, the r^2 between behavior and any individual feature is smaller than < 0.16 (all r^2 not cross-validated). This indicates that the stimuli are not linearly separable in the Fourier domain.

Recently, a neural network was trained successfully on the same data set, yielding a performance of $r^2 = 0.83$ (Wittmann et al., 2011). The network based its prediction on 7 temporal features, proposed in Schmidt et al. (2008) and based on previous behavioral experiments. I performed simple multi-linear regression on these 7 features, which yielded a performance comparable to that of the highly nonlinear network ($r^2 = 0.79$, not cross-validated). Regression against individual features yielded two temporal features—"pronounced end" and "plateau length"—with moderate predictive performance of 0.45 and 0.53 (not cross-validated), respectively.

The "naive" approach applied here yielded a single feature that outperformed the best individual feature used in Wittmann et al. (2011) by $\approx 20\,\%$ ($r^2 = 0.64$, cross-validated, Fig. 4.4 c, g). Note that the approach achieves this performance without incorporating prior knowledge on the relevance of specific features on signal attractiveness—the approach yields thus the feature *and* the classifier.

As for the block-like stimuli, I trained a single classifier to the full data set, without performing the cross-validation. Performance was not higher than the cross-validated one indicating that over-fitting was no problem. Additionally, classifier structure resembled that obtained by cross-validation.

Interestingly, the filter resembled that obtained for the block-like stimuli in that it was also an offset detector with a leading positive and a lagging negative lobe (Fig. 4.11 a, compare Fig. 4.7 a). However, the leading positive lobe appeared to be broader. Additionally, the nonlinearity was much shallower with a dynamic range covering almost the full range of filter outputs (Fig. 4.11 b); the output of this nonlinearity resembled thus much more strongly the filtered stimulus than the output of the classifier for the block-like stimuli.

The classifier performed well for most stimuli but exhibited a few outliers Fig. 4.11 c): overall performance was 0.64, excluding the five largest outliers increased performance to 0.77, indicating that the classifier performed well for the majority of stimuli. In order to

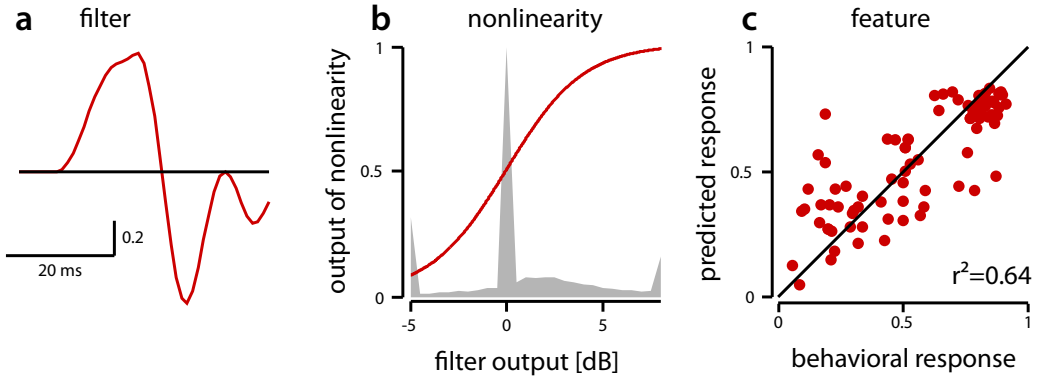

Figure 4.11.: **Structure of the classifier for natural-like stimuli.** **a** linear filter, **b** nonlinearity transforming the filter output. Grey area depicts the distribution of filter outputs. **a** Measured vs. predicted response. Each dot corresponds to one stimulus. The predicted response is obtained by linear regression of the average output of the nonlinearity against the measured responses.

understand how the classifier works, let us look at the five best and worst predicted stimuli (Fig. 4.12).

Stimuli with the smallest prediction error Among the best-predicted stimuli were those resembling a "standard syllable" (Fig. 4.12 d 7 and 8) with an accentuated onset and a slowly decreasing syllable plateau. Other stimuli with minimal prediction error were those with ragged or interrupted plateaus (Fig. 4.12 d 6 and 10), for which attractiveness was low. Furthermore, a shallowly sloped syllable onset was rejected both by the animals and the classifier as well (Fig. 4.12 d 9). The predictions thus conformed to the intuition expressed in Schmidt et al. (2008) that smooth syllables are attractive while interrupted ones are unattractive. A look at the output of the filter-nonlinearity cascade provides insights into why these intuitions hold: Remember that the filter preferred offsets or negative slopes in the stimulus while the nonlinearity was relatively shallow. Hence, patterns with a single onset and an uninterrupted, slowly decaying plateau produced large outputs (Fig. 4.12 e 7 and 8). In contrast, the output for stimuli with an interrupted onset were suppressed during that epoch, which introduced large pauses in the output and hence small overall feature values (Fig. 4.12 e 9). The same applied to those stimuli with interrupted plateaus (Fig. 4.12 e 6 and 10) as these also contained many, "suppressive" onsets and hence produced weak outputs.

Stimuli with the largest prediction error The best-predicted stimuli thus show that the classifier reproduced the behavioral sensitivity to ragged onsets or plateaus. There were, however, instances were the classifier weighted them inappropriately. Two behaviorally attractive stimuli were predicted as being only moderately attractive (Fig. 4.12 d 3 and 4, 0.87 vs 0.48 and 0.79 vs 0.43). These stimuli exhibited a dip in the plateau shortly after the onset, which introduceed another shallow onset and thereby suppressed output during the syllable (Fig. 4.12 e 3 and 4). Probably, such dips shortly after a loud stimulus epoch are not detected by the animal (Stefanie Krämer, unpublished data). Two stimuli with an inter-

rupted syllable end and a ragged plateau evoked low behavioral scores yet were predicted as being attractive (Fig. 4.12 d 1 and 5, 0.19 vs 0.73 and 0.19 vs 0.54). In the first case (stimulus #1), this probably was because of the long slow offset producing a large output at the syllable end (Fig. 4.12 e 1). In the second case (stimulus #5), a dip in the plateau in the middle of the syllable strongly suppressed behavior while the filter output was only weakly suppressed (Fig. 4.12 e 5). Compare this stimulus #5 to two stimuli #3 and #4 which were predicted with comparable attractiveness yet evoked much higher behavioral scores. They differed mostly in the timing of the dip relative to the onset. While dips at the syllable beginning did not affect behavior, they did so if at the end of the plateau. The classifier failed to make this discrimination and rejected all stimuli with a dip. Using a longer filter or incorporating an adaptive pre-processing step (see e.g. Creutzig et al. (2010)) might solve this problem. Adaptation would suppress firing shortly after the onset and make the classifier thereby insensitive to "early" dips in amplitude. The other stimulus for which attractiveness was overestimated was stimulus #2 with a slowly rising and smooth syllable plateau (Fig. 4.12 d 2, 0.16 measured vs 0.57 predicted). The output was not unlike that of the well-predicted and attractive stimuli #7 or #8. This failure to predict the low attractiveness of this stimulus need not necessarily reflect a weakness of the classifier. Rather it could be that the features underlying the behavioral rejection of such stimuli were rare in the stimulus set for the training algorithm and could thus not be learned effectively.

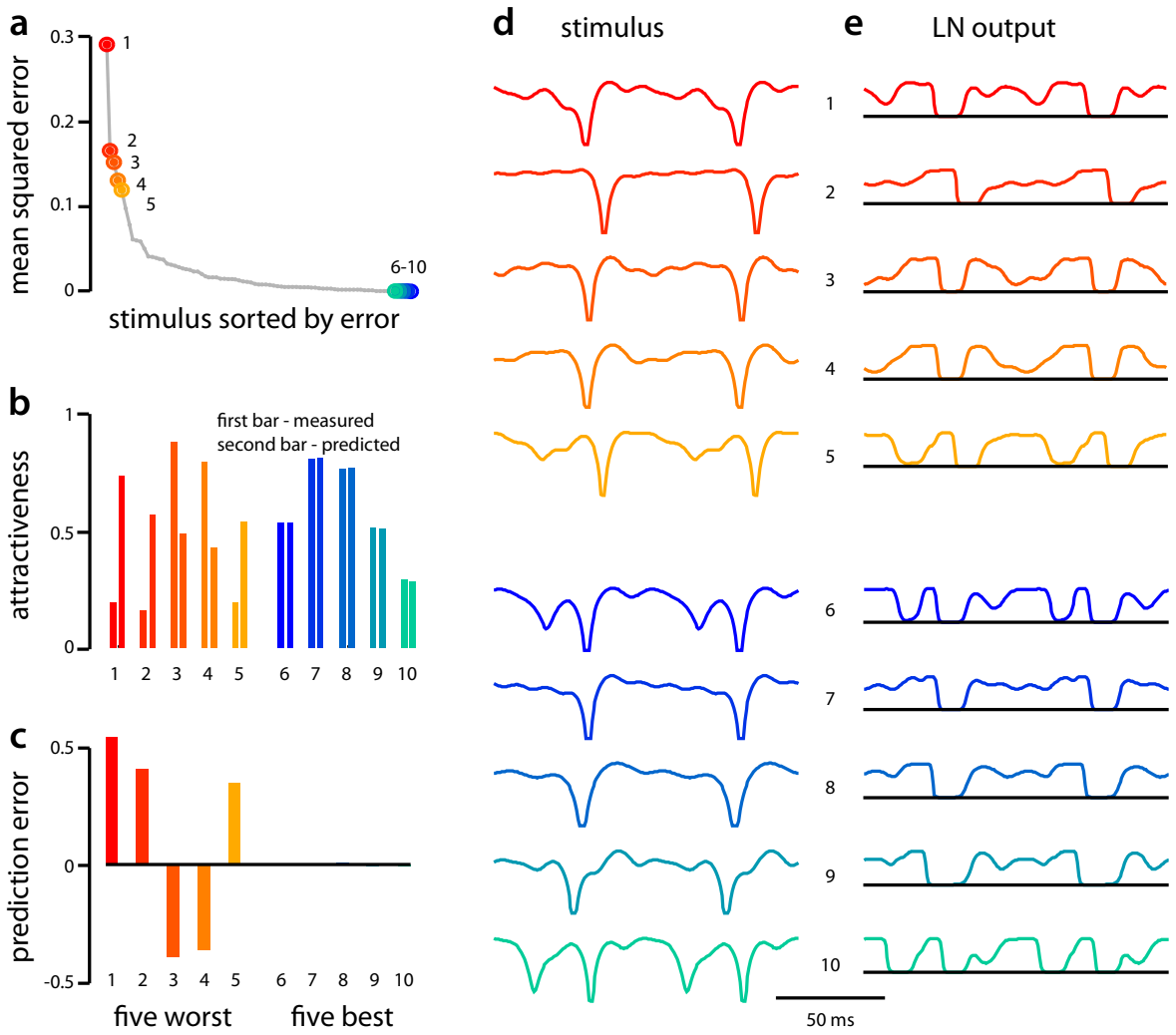

Figure 4.12.: **Natural-like stimuli with the largest and smallest prediction error.** The five patterns with the largest prediction errors are marked with large reddish dots, those with the smallest prediction errors with blueish dots. **a** Mean squared errors of all stimuli sorted in descending order. **b** Behavioral and predicted attractiveness. The height of the first bar in each pair indicates measured, that of the second predicted attractiveness. **c** Difference between behavioral and predicted attractiveness (height first bar - height second bar). This corresponds to the *non*-squared error. Bars for the five best stimuli are not visible because the error was essentially zero. **d** Envelope of the five best and worst stimuli for which the error is largest (dB scale). **e** Output of the filter and nonlinearity for the five best and worst stimuli in (a). Black vertical lines indicate zero output.

4.4. Discussion

This chapter attempted to test the hypothesis that the timing of temporal features of grasshopper song is unimportant for song recognition as has been suggested by the properties of the early auditory system of grasshoppers (chapter 2). The results have shown that a simple classifier implementing this hypothesis performs well on block-like and natural-like stimuli. Furthermore, the structure of the learned classifier provided insights into how different features of a stimulus influence behavior. In the following, I will discuss these points further, ask whether the "when" of temporal features is necessary to explain song recognition in grasshoppers and propose extensions which might further improve performance.

4.4.1. Can the approach inform hypotheses about the neural implementation of song recognition in grasshoppers?

The structure of the classifier was inspired by the early auditory system of grasshoppers. The stages of the classifier can be identified with different processing steps in the early auditory system of grasshoppers: auditory receptors as encoders of the envelope, ascending neurons as feature extractors and an feature integrator in the brain. This relatively simple structure was capable of explaining many aspects of song recognition in grasshoppers (Fig's 4.5, 4.6, 4.8 and 4.10).

To what extend can it inform hypotheses of the neural implementation of song recognition? First and foremost, the classifier for block-like stimuli yielded a fundamental and unexpected insight: studies looking for neural correlates of behavioral selectivity usually do so by searching for a high correlation between neural and behavioral responses. However, I found that behavior can be well explained by the linear combination of two features that are in themselves only poorly correlated with behavior (Fig 4.7 d, e). Hence, population-level approaches are more likely to be successful in relating neural coding and behavior if single-neuron responses are only poorly correlated with behavior (compare the multi-neuron metric applied in chapter 2).

The linear filters optimally predicting female responses for block-like and natural-like stimuli were offset detectors. This is somewhat at odds with previous hypotheses about song recognition in grasshoppers: Behavioral studies have shown that accentuated syllable ends are unattractive (von Helversen and von Helversen, 1987). Along this line, electrophysiological studies have shown that the ascending interneuron AN12 encodes the pause duration for block- and natural songs. AN12, however, detects the *onsets* of a syllable and produces responses that are proportional to the duration of the preceding pause through an interplay of adaptation and slow inhibition (Creutzig et al., 2009). Furthermore, the analysis of computations underlying sparseness in the auditory system grasshoppers (chapter 3) has shown that onset—and not offset—detection is a basic computation implemented in many ascending neurons.

However, the classifier is defined on an algorithmic or computational level, not on a mechanistic level. Hence, many implementations could yield equivalent computations. Indeed, the offset detector can easily be transformed into an onset detector by a change of sign:

putting a negative sign in front of the filter would yield an onset detector. Flipping the x-axis of the nonlinearity would create an onset detector with a similar output than the offset detector.

This shows that one cannot directly translate the structure of the classifier model to a neural implementation. Hence, prior knowledge about the physiology of a system needs to be taken into account to constrain possible implementations of a computation. This could be accomplished directly by using the LN models fitted to auditory neurons in chapter 3 as a feature detection stage.

While the shape of the filter(s) does not conform with what one would expect, the tuning for pause duration of the feature detectors obtained for the block-like stimuli does: recent, unpublished results by Olaf Kutzki suggest that a pair of ascending neurons, AN3 and AN12 exhibit a pause tuning similar to that found for the two feature detectors for block-like stimuli (Fig. 4.8): AN3 is a weak band-pass filter with a slowly falling flank for long pause durations while AN12 is a pure high-pass with a nearly monotonous tuning. Both neurons are usually described as detectors of onset (Creutzig et al. (2009), chapter 3). This further suggests that the feature detector is implemented as an onset rather than an offset detector.

Interestingly, one or two feature detectors were sufficient to predict behavioral responses to the stimuli in the data sets. Yet, the auditory system of grasshoppers contains 15–20 ascending neurons, the neurons that likely correspond to the feature detectors in the classifier. However, the data sets used here did not cover the whole range of signals and functions relevant for the auditory system of grasshoppers. The stimuli only covered a small range of the naturally occurring stimulus patterns that are usually rejected by female *Chorthippus biguttulus*. More than two feature detectors might be needed to accomplish this robustly. Additionally, intensity invariance at the level of ascending neurons appears to be relatively low, despite behavior working over a wide range of intensities (Weschke and Ronacher, 2008; Clemens et al., 2010). The diversity of ascending neurons could serve to establish intensity invariance at a later stage, by computations integrating the responses of different ascending neurons (see Römer and Seikowski (1985) and Uchida and Mainen (2008) for an example in olfactory coding). Additionally, although the songs differ greatly between species of the *Chorthippus* group, the selectivity of ascending neurons seems to be conserved within this taxon (Neuhofer et al., 2008). The large diversity of neurons might underly the diversity of behavioral selectivities and the rapid radiation of species in the group (Bugrov et al., 2005). Moreover, the auditory system of grasshoppers also serves the recognition of predators and the localization of sound sources (Robert, 1989; Schul et al., 1999; von Helversen, 1997). There exist ascending neurons which are especially sensitive to bat echolocation signals and which encode the direction of a sound (Stumpner, 1988; Stumpner and Ronacher, 1994).

4.4.2. No timing required?

The classifier was built to test the hypothesis that song recognition in grasshoppers is sufficiently explained by knowledge about the "what" of temporal features and that their "when" is not necessary (see chapter 2). This was motivated by the finding that the brain of grasshoppers receives only imprecise information about the timing of features but very spe-

cific information about the nature of these features. This last property was correlated with an increased sparseness and a highly diverse selectivity for temporal features of the sound at the output of the metathoracic ganglion. As the hypothesis entails a very simple mode of pattern classification, this would support the notion that population sparseness facilitates subsequent computations (Olshausen and Field, 2004). More specifically, the population sparseness at the level of ascending neurons would facilitate song recognition in grasshoppers by allowing the brain to ignore the temporal structure of the responses of ascending neurons and simply count their spikes.

One pre-requisite for this hypothesis being true is that the classifier successfully predicts behavioral responses—in the majority of the stimulus sample, this was indeed true. The performance of the classifier was very good for block-like stimuli. As this stimulus set probed the perception of pause duration and its dependence on specific features surrounding the pause like the offset, onset and intensity (Fig's 4.4 a,b), I can conclude that global timing of pauses is not needed to reproduce the behavioral evaluation of this feature in this stimulus set. This is corroborated by behavioral experiments which have shown that syllable-pause pattern of a song can be scrambled without affecting signal attractiveness as long as the the average duration of the pauses throughout the songs is optimal (von Helversen and von Helversen, 1998).

its performance for natural stimuli was limited by small variance in the behavioral data (Fig. 4.4 d). Furthermore, to show that the classifier works with natural songs as well, extension of the data set with the songs of other members of the *Chorthippus* group should

The classifier also performed well for natural-like stimuli with some outliers (Fig's 4.4 c, 4.12). Hence, even for the more richly modulated natural-like stimuli, timing could be largely ignored when accounting for song recognition. Does the presence of outliers argue against the hypothesis that timing is not necessary to explain behavior? Not conclusively. The classifier relies on many assumptions. The classifier's failure for some stimuli indicates that one of the assumptions does not apply to the song recognition system of grasshoppers—not necessarily that information about the timing of features is required.

1. First, **adaptation** has been shown to act at all stages of auditory processing in the grasshopper (Benda et al., 2001; Hildebrandt et al., 2009). However, my model does not include adaptation, an omission that might account for some failures of the classifier for natural-like stimuli. Some outliers in this data set could be explained by the classifier not being able to discriminate between a dip shortly after the onset and at the end of the plateau (Fig. 4.12). The "positional" differentiation of the impact of a dip depending on its timing relative to the syllable onset could well be provided through a longer filter or spike frequency adaptation at some stage in the classifier (compare Creutzig et al. (2010); Pillow et al. (2008)).

2. Second, the classifier is **linear**—both at the initial filtering stage as well as the transformation of feature values to the behavioral responses (regression step). The assumption of linearity in the feature detection stage is likely to by justified, as neurons at all processing stages in the early auditory system of grasshoppers seem to be well described by linear-nonlinear models (Machens et al. (2001), chapter 3). Using the two-

dimensional models for ascending neurons obtained using spike-triggered covariance analysis in chapter 3 could further improve classification success. In contrast, the final regression step is likely to be *non*-linear. Incorporating an explicit nonlinearity in the regression step might increase performance(Jäkel et al., 2009).

3. The third assumption is that the **data is sufficient** to train the classifier. While this was true for the very simple block-like stimuli, where classification performance was very good, some outliers for natural-like stimuli (Fig. 4.12) can probably be explained by insufficient data or inefficient training. These stimuli were more richly modulated than the block-like stimuli, which might constitute "noise" the classifier first has to learn to ignore. Along this line, the features underlying the behavioral evaluation of the outliers might have been too rare in the data set to be picked up during the training. This problem could be solved either by enlarging the data base or by optimizing training. The latter solution could be implemented by applying more robust regression schemes, imposing further constraints on the parameters, either informed by knowledge about the system (allow onset-shaped filters only) or by statistical learning theory (e.g. impose sparseness on filter weights, see Olshausen and Field (1996); Mineault et al. (2009))

Can one say that grasshoppers do indeed ignore the global timing of features when evaluating a song? The results obtained on different data sets suggests that the "when" of temporal features is *not* necessary to explain most behavioral data. Whether behavior can be fully explained using a model that solely relies on the "what" of features of the song needs more natural-like data and a more efficient classifier.

4.4.3. Do the results provide evidence for a role of population sparseness in facilitating song recognition?

The results show that many aspects of song recognition in grasshoppers can be explained by a classifier using only the average output of features over a song. This implements a simple mode of perceptual decision making in a shallow neural network. It is simple, because decision making centers in the brain do not have to evaluate the temporal pattern of neural inputs but only the presence/absence or the number of spikes. I expected that a large specificity and diversity of features would be necessary to enable this simple mode (see chapter 2).

Surprisingly, no more than two, relatively similar features were sufficient to maximize classification performance (Fig's 4.7 and 4.11). Hence, the classifier did not rely on a large diversity of different features but rather on a *single* feature associated with different nonlinearities. It did thus not leverage population sparseness.

Note, that although the classifier was not *population* sparse, the output of the features appeared to be *temporally* sparse (Fig. 4.9). This replication of a response feature of ascending neurons by the classifier is an emergent property of the classifier. This implies a high selectivity of the feature detectors—they respond only if "their" feature is strong. This increases the signal to noise ratio of the feature values and might be necessary when relying

on time-averaged features.

Hence, while the classifier shows that a simple mode of song recognition is likely to be implemented in the grasshopper, it does not prove that population sparseness enables this simple mode. To directly show that the song recognition system of grasshoppers is an example for how sparse coding can facilitate subsequent computations, additional studies are necessary. Increasing the diversity of stimuli and improving the classifier as proposed above might reveal the necessity for a larger diversity of stimulus features—and hence population sparseness—to be able to explain behavior without considering the timing of stimulus features.

4.4.4. Conclusion

The classifier has given useful insights into how the envelope of a song influences female responses. Furthermore, it has advanced our understanding of how different features could be evaluated and integrated by decision making centers in the brain of grasshoppers. The approach can also be used to investigate perceptual decision making in other species or sensory modalities. E.g. in visual psychophysics, the classifier could be modified to become invariant to spatial translation, by averaging the output of a filter at different locations. It could also be applied to compare song recognition in closely related species (Stumpner and Helversen, 1992) or in male and female individuals (von Helversen and von Helversen, 1997; Neuhofer et al., 2011), or to explain the selectivity of hybrids between two species (Gottsberger and Mayer, 2007; Vedenina et al., 2007). It might even be useful to estimate neuronal filters in the case were spike-timing is very loose, e.g. in a rate-based coding scheme. The classifier might also prove useful to test normative theories of coding (see e.g. Geisler et al. (2009); Olshausen and Field (1996); Smith and Lewicki (2006)): Given a set of natural stimuli and a naturally occurring task (recognition of conspecific song), one could train the classifier to accomplish this task and compare its structure to that of the biological system fulfilling the same job.

Part II.

Adaptive coding in the auditory system of the cricket

5. Stimulus-dependent spectrotemporal tuning in the auditory system of the cricket: mechanisms and consequences for information transmission

In this chapter, I will change both the model system as well as the question. I will start from basic considerations of how a neural code should change with changing stimulus complexity. I will then explore whether the auditory system of the cricket employs stimulus-dependent codes and propose putative mechanisms implementing these.

5.1. Introduction

Natural sensory environments are composed of many different signals, each of which is a complex mixture of several elements in itself. Sensory systems can benefit from the co-occurrence of several stimulus components, since ambiguities can be resolved and invariant object recognition achieved (Uchida and Mainen, 2008; Riffell et al., 2009).

However, the presence of multiple stimulus elements can also be detrimental to object recognition by masking and distorting important information (Narayan et al., 2007; Schmidt et al., 2011). Especially at higher stimulus intensities, receptor tuning often becomes broad and thus un-selective (Hallem and Carlson, 2006). Then, different stimulus components interfere with each other in the response of a single neuron. Such a degradation of the neuronal representation of individual components will introduce ambiguities for downstream neurons.

It comes to no surprise that sensory systems are adapted to cope with multi-component stimuli. A common solution for the interference problem is to employ adaptive coding by context- and stimulus-dependent codes (Vinje and Gallant, 2000; Schneider and Woolley, 2011; Ahrens et al., 2008; Geffen et al., 2009; Chacron et al., 2005; Machens et al., 2004). Such phenomena are generally interpreted to optimize coding either by exploiting valuable context information in non-linear combinatorial codes (e.g. responses to bird's own song (Margoliash and Konishi, 1985)) or by suppressing background noise.

The role of inhibition for central sensory coding is very much under debate at the moment (Hasenstaub and Callaway, 2010; Schneider and Woolley, 2011; Isaacson and Scanziani, 2011). While the physiology of inhibitory cells is well understood, its function *in vivo* is hard to elucidate (Kerlin et al. (2010); Runyan et al. (2010), but see Olsen and Wilson (2008)). Inspired by a small model system with only three cells and a well-defined task—the auditory system of the cricket—a modeling approach will show how inhibition helps to solve the interference problem and improves the coding of complex stimuli. Recordings from the auditory system of the cricket will be employed to support these findings.

The cricket's acoustic world is divided into two frequency ranges associated with different ecological meaning (Wyttenbach et al., 1996). Low carrier frequencies are associated with mating signals and elicit approaching behavior. High carrier frequencies induce avoidance behavior as they are associated with the echolocation signals of cricket-hunting bats.

This dichotomy in the semantics of sensory stimuli is reflected in the organization of the early auditory system of the cricket: Two populations of receptors, one most sensitive to the carrier frequency of courtship signals, one being tuned to ultrasound of bats, form the input to a small, three-neuron network in the animal's prothoracic ganglion (Kostarakos et al., 2009; Imaizumi and Pollack, 1999). The network has two output neurons (ascending neurons – AN) which project to the brain for further processing of the stimulus. Ascending neuron 1 (AN1) encodes the amplitude modulations at low carrier frequencies (Schildberger et al., 1988; Hennig, 1988). Ascending neuron 2 (AN2) encodes ultrasound signals and is necessary and sufficient for negative phonotaxis (Nolen and Hoy, 1984; Marsat and Pollack, 2006). Both AN are inhibited by a single, broadly-tuned omega neuron 1 (ON1) (Selverston et al., 1985).

Although the seemingly simple structure of the cricket's auditory world suggests sharp tuning and a clear separation of both frequency ranges, there are several sources of interference between both channels in the system. First, receptors are relatively broadly tuned—at high intensities, receptors will respond to all carrier frequencies (Hennig et al., 2004; Imaizumi and Pollack, 1999). Second, some but not all neurons pool input from both receptor populations, further degrading the separation of both frequency ranges (Pollack, 1994; Nolen and Hoy, 1987).

An abstract, linear encoding model suggests that the interference problem can be solved by sharpening the tuning at the encoder side (Schmidt et al., 2011). A simple network model and experimental data from the early auditory system of the cricket show that this sharpening can be easily accomplished in an adaptive and stimulus-dependent manner with a static network and two mechanisms that are found in many nervous systems: a logarithmic input nonlinearity in the periphery and a broadly-tuned feed-forward inhibition.

5.2. Methods

5.2.1. Electrophysiology

Animals: Adult female *Gryllus bimaculatus* were obtained from a commercial supplier and kept isolated from males in a 24 h light-dark cycle. After visually inspecting the intactness of the tympana, the mid and hind legs as well as the wings were removed and the animal was dorsally fixed to the recording stage with wax. The front legs were fixed in a roughly natural position while care was taken not to restrain the tympana with wax. In order to reduce upstream neuronal activity and body movements, both the meso- and the metathoracic ganglion were removed. Maxillae, labrum and gut were removed as well.

Recording: Recordings were performed by Florian Rau. Signals from AN1 and AN2 were recorded differentially from one of the connectives between prothoracic and suboesophageal ganglion using tungsten hook electrodes, referenced to a silver wire in the animal's abdomen (Hennig, 1988). Signals from ON1 were recorded in separate sessions from the prothoracic ganglion using an extracellular tungsten electrode (World Precision Instruments, Sarasota, FL, USA), referenced to a stabilizing metal spoon.

Voltage signals were bandpass-filtered between 300 and 3000 Hz (DPA-2FX, npi electronics, Tamm, Germany), digitized at 20 kHz sampling rate (PCI-6229, National Instruments, Austin, TX, USA) and recorded on a PC using LabView software (National Instruments).

After digital-to-analog conversion, audio stimuli were adjusted to the desired sound pressure level with an attenuator (ATN-01M, npi electronics), amplified with a power amplifier (Raveland XA-600, Blaupunkt, Hildesheim, Germany) and presented via one of two loudspeakers mounted on either side of the animal. Sound intensity was calibrated (1/2 inch microphone, type 2209, Bruel and Kjaer) using pure tones with the carrier frequencies used in the experiments (4.5, 10, 15, 25 kHz) at an intensity of 80 dB SPL as well as a sum of all four carriers.

Spike sorting: Action potentials were detected by a threshold in the differentiated voltage traces. If SNR allowed it, both AN were detected in the same voltage trace and sorted by the amplitude of their spikes as AN1 has smaller amplitude spikes than AN2 (Hennig, 1988). In addition, the identity of the recorded cells was ensured by their physiological characteristics (sensitivity to different carrier frequencies, maximal firing rates, sensitivity to contralateral input, Marsat and Pollack (2004)). The data set consists of 5 specimen of ON1, 6 AN1 and 14 AN2.

Stimuli: Stimuli consisted of pure tones with a frequency of 4.5, 10, 15, or 25 kHz, whose amplitude was randomly modulated by Gaussian low-pass noise with a cutoff of 200 Hz, a mean of 80 dB SPL and a standard deviation of 6 dB. 4.5 kHz corresponds to the carrier frequency of the male song, 15 and 25 kHz are associated with the echolocation signals of bats and 10 kHz is intermediate between these two frequency ranges. These carrier frequencies were represented either alone ("single carrier") or as a sum ("composite carrier"). Note,

that in the composite case, the amplitude of each one of the four carriers was modulated with independent noise—different bands were thus uncorrelated.

For estimating the linear-nonlinear models 20 different stimuli were presented, each of which lasted 20 s. The time-varying firing rate for evaluating the models and computing information was estimated from 20–40 repetitions of another 4 s stimulus. As all three neurons adapted, we used only the stationary part of response, omitting the first 0.4 s of each spike train (see e.g. Benda and Hennig (2008)).

5.2.2. Estimation of linear-nonlinear models.

From the responses to the long noise stimuli, we estimated linear-nonlinear models. Such models consist of a linear filter which describes a cell's selectivity for (spectro-) temporal features of the stimulus and a nonlinearity, which relates the filter's output to the cell's firing rate and depicts the tuning of the cell for the filter (see chapter 3, Schwartz et al. (2006)).

Filters were estimated as spike-triggered averages (STA). To that end, the single-carrier stimulus was down-sampled to 1000 Hz and the average envelope in the 64 ms preceding a spike was calculated. Doing this for all four carrier frequencies yielded a set of four filters. For the composite stimulus, consisting of four frequency bands the amplitude of each of which was modulated independently, the spike-triggered average was calculated for each carrier frequency separately, yielding another set of four filters. This corresponds to a low-resolution spectro-temporal receptive field (STRF), used to analyze the joint-tuning of cells for spectral and temporal features of a stimulus (see e.g. Atencio et al. (2008)).

The nonlinearity was estimated by filtering the stimulus with the normalized STA and applying Bayes' rule: $p(r|s) = p(r|s)/p(s) \cdot \langle r \rangle$ (for details see chapter 3 and Schwartz et al. (2006)). $p(s)$ is the amplitude distribution of the filtered stimulus, $p(s|r)$ is the amplitude distribution of the filtered stimuli preceding each spike and $\langle r \rangle$ is the cell's average firing rate. I fit a Gaussian to both $p(s)$ (which was by definition Gaussian) and $p(s|r)$ and calculated the nonlinearity as a ratio of two Gaussians scaled by the average firing rate (Pillow and Simoncelli, 2006). This parametric approach did not reduce the model performance when compared to a non-parametric approach but gave smoother nonlinearities especially for poorly sampled extreme values of the filter output.

For the composite stimulus, the nonlinearity for each frequency band was estimated separately. Each neuron was thus modeled as a bank of four parallel filters and nonlinearities, one for each carrier. The four inputs were added to form the firing rate of each neuron. Note, that this approach is different from the standard STRF-model, which has a single nonlinearity for the pooled output of each carrier frequency's filter.

To calculate the average nonlinearity for a given cell type, carrier frequency and stimulus condition (simple or composite carrier), I divided the nonlinearity by the average firing rate $\langle r \rangle$ prior to averaging individual cells and then rescaled the average nonlinearity by the mean average firing rate over all specimen.

Model performance was quantified on a novel stimulus not used for the estimation of the filters and nonlinearities. The time varying firing rate of the neuron was estimated from multiple representations of a short noise segment, by binning the spike trains with a res-

olution of 1 ms and smoothing the resulting time-varying firing rate with a box window spanning two bins. To predict the response, the stimulus envelope was down-sampled to 1000 Hz and fed into the linear-nonlinear model. A bias-corrected coefficient of correlation between the actual and the predicted response served as a measure of model performance (see chapter 3, Petersen et al. (2008)). Predictive power of the models was 0.65 ± 0.16 for all cells and all stimulus conditions indicating that LN models were appropriate for describing the computations performed by the cells in our data set.

The distance d between filters for single and composite stimuli was calculated as one minus the dot product between normalized filters for each cell type and carrier frequency: $d(f_1, f_2) = 1 - f_1^T f_2$, where f_i are the unit-norm filters as column vectors and ^{T} is the transpose operator.

As the most prominent change of the nonlinearity with stimulus condition was a reduction of the slope, we estimated a gain factor for the average full nonlinearity for each cell type, by calculating the average slope of the nonlinearity.

5.2.3. Abstract encoding model and coherence information

The model encoder received Gaussian, uncorrelated input on two channels; the encoder filtered and thresholded the stimulus in both channels; then, both channels were pooled to yield the output firing rate of the cell. Filters were taken from a recorded AN2 at 25 kHz from the composite condition, consisting of a biphasic filter at 4.5 kHz and a low-pass filter at 25 kHz (Fig. 5.5 b, filter pair in lower left panel). The impact of the decorrelation of filters and the suppression of output gain was examined by calculating the coherence information in the pooled firing rate about the envelope in each individual input for different forms and scales of the filters. Note, that as the nonlinearity was fixed with a slope of 1 Hz/dB, reducing the filter's scale is equivalent to reducing the cell's output gain. Coherence information estimates the information retrievable by optimal linear reconstruction and yields a lower bound on the mutual information between stimulus and response (Borst and Theunissen, 1999; Roddey et al., 2000). It is given by $I(f) = -\log_2(1 - C_{sr}(f))$. Stimulus-response coherence is defined as $C_{sr}(f) = |P_{sr}(f)|^2 / P_{ss}(f) / P_{rr}(f)$, where P_{sr} is the cross-spectral density of the stimulus envelope s and the response r, and P_{ss} and P_{rr} are the respective autospectral densities. The total information was obtained by integrating $I(f)$ for each input between 0 and 100 Hz and summing the information for both inputs. Calculation of information in the responses of electrophysiological recordings of ON1, AN1 and AN2 was performed as for the model and was based on the time-varying firing rates estimated as described for the model evaluation.

5.2.4. Network model

A minimal network model inspired by the auditory system of the cricket was built to examine how stimulus-dependent coding can be implemented. This minimal circuit consisted of two receptor populations (low- and high-frequency), ON1, and the two ascending neurons AN1 and AN2. Stimulus was provided by four input channels corresponding to the

four carrier frequencies (Fig. 5.2 a). Frequency tuning of the receptor populations was modeled by frequency-specific attenuation (Fig. 5.2 b): Input to low-frequency receptors was most strongly attenuated at 15 and 25 kHz; effective input to high-frequency receptors was weakest at 4.5 kHz. This was implemented as a gain factor with which each carrier frequency was multiplied. The firing rate of each receptor population corresponded thus to a weighted sum of the envelopes of the four carrier frequencies on a pressure scale that was then transformed to a dB scale by a logarithmic nonlinearity (Fig. 5.2 c). To have the time scale of the filters in the model match those seen in the experimental data, a low-pass filter was applied to each driving current (Gaussian, σ =7 ms).

Synapses were modeled with a weight (positive for excitatory, negative weight for inhibitory inputs) and a delay. The time-varying firing rates of the three neurons in the network were thus the weighted and delayed sum of their inputs (Fig. 5.2 d). A list of all model parameters (tuning and nonlinearity of receptors, synaptic weights and delays for ON1, AN1 and AN2) can be found in table 5.1.

To analyze the encoding properties of the cells in the model, the same framework of linear-nonlinear models as in the analysis of the experimental data was employed. Filters were calculated as the cross-correlation between the stimulus envelope and the modeled firing rate. Due to a lack of a spiking nonlinearity or threshold, the "static nonlinearity" is linear with a gain/slope corresponding to the norm of the filter. Note, that as the network is linear, the gain is only determined up to a constant for each cell.

parameter name	value
attenuation of receptors	**at 4.5, 10, 15, and 25 kHz**
low frequency, LF	0, 8, 16, and 16 dB
high frequency, HF	16, 8, 0, and 0 dB
synapse	**delay, weight**
LF to ON1	10 ms, 0.4
HF to ON1	6 ms, 0.4
LF to AN1	12 ms, 0.3
LF to AN2	12 ms, 0.1
HF to AN2	10 ms, 0.6
ON1 to AN1	6 ms, -0.15
ON1 to AN2	5 ms, -0.35

Table 5.1.: **Parameters of the network model**

Receptors Population responses of LF and HF receptors were approximately linear within the dynamic range considered here (Imaizumi and Pollack (2001), Ulrike Ziehm, unpublished results). Consistent with a subtractive effect of the frequency tuning of receptors, the tuning curves of individual receptors were identical up to a translation along the dB axis, while the slope was constant (Imaizumi and Pollack, 2001; Gollisch et al., 2002).

ON1 Received input from low- and high-frequency receptors (Hirtz and Wiese, 1997). Inputs from low-frequency receptors were slower than those from high-frequency receptors (Pollack, 1994; Faulkes and Pollack, 2000). Since inclusion of contralateral attenuation had no effect on filters or gain, it was not implemented in the model.

AN1 Receives excitatory input from low-frequency receptors only (Hennig, 1988; Hirtz and Wiese, 1997). Inhibitory input from ON1 was coincident with this low-frequency input (Hardt and Watson, 1994; Selverston et al., 1985; Faulkes and Pollack, 2000; Wohlers and Huber, 1982).

AN2 Received excitatory input from low- and high-frequency receptors (Wohlers and Huber, 1982; Hennig, 1988; Nolen and Hoy, 1987; Hirtz and Wiese, 1997). Input from low-frequency receptors and ON1 was slower than that of high-frequency receptors (Hardt and Watson, 1994; Pollack, 1994).

5.2.5. Statistics

Data analysis and modeling was performed with custom-written routines in Matlab. Results are reported as mean$\pm$s.e.m over specimen. For multiple comparisons (across carrier frequencies) a Kruskal-Wallis test with a Tukey-Kramer post hoc test was used (Matlab's `multcompare`), for pair-wise comparisons (single- vs. composite carrier stimuli) a sign test was applied.

5.3. Results

5.3.1. Decorrelation and suppression preserve information when encoding multiple stimuli

To gain a more rigorous understanding of the interference problem—and to conceive possible solutions to it—I implemented a simple encoder which filters and thresholds two different, time-varying inputs and pools them in a single output channel. In the cricket, this would correspond to a single neuron encoding the envelope of two different carrier frequencies. I then asked how well an optimal, linear decoder can reconstruct both stimulus waveforms from the pooled output and searched for the optimal filter prior to the pooling to maximize the information transmission of such a system. This problem is akin to the multiple-access channel known in information theory (Cover and Thomas, 1991; Fischer and Westover, 2003). One solution turned out to be frequency multiplexing, a method that has found technical application in cable transmission. Here, I present these results in an intuitive fashion and apply it to sensory coding.

To get a baseline value for the information this system can transmit without interference, I asked how well a single, isolated input can be decoded, obtaining a value of 110 bit/s (Fig. 5.1 b). This value was greatly reduced to 80 bit/s when pooling two inputs filtered with identical filters (Fig. 5.1 c). Note that this represented the information about both inputs. Information about each individual component was only 40 bit/s although the power of *each* input matched that of the single input in Fig. 5.1 b and hence total input power was doubled. The main problem when reconstructing two input components from a single, pooled output is that of ambiguity: given only the pooled firing rate, it is not clear which response components correspond to which input, as the spikes in a pooled response lack a label marking them as coming from either input (compare chapter 2 and Houghton and Sen (2008); Clemens et al. (2011)). Even worse, for reconstructing one of the two inputs, the response component corresponding to the other input constitutes noise, thereby reducing the signal-to-noise ratio and inducing a loss of total information. Interference between both inputs can be reduced by minimizing the correlation between the response components associated with each input. In a linear system, this can be achieved either by changing the shape of the filters or their gain.

Changing filter shapes such that they are decorrelated equipped them with different pass bands, resulting in a filter pair with e.g. a low and a high pass band (Fig's 5.1 d). That way, different frequency bands in the response were preserved for representing the corresponding frequency band of one of the two stimuli. In the case of a low- and a high-pass filter, one input was encoded in low-frequency components of the response, while the other input was encoded in high-frequency components. Given no overlap between both pass-bands, the decoder could now unambiguously decode the encoded frequency range of each input by applying the appropriate decoding filter. However, such frequency multiplexing comes at a drawback: only the part of each input that was let-through by the respective filter was represented in the response. Thus, while the decorrelation of the filters for both inputs increased total information, it still led to a loss of information about parts of each input. This is not detrimental if relevant information occurs at different frequency bands for each input.

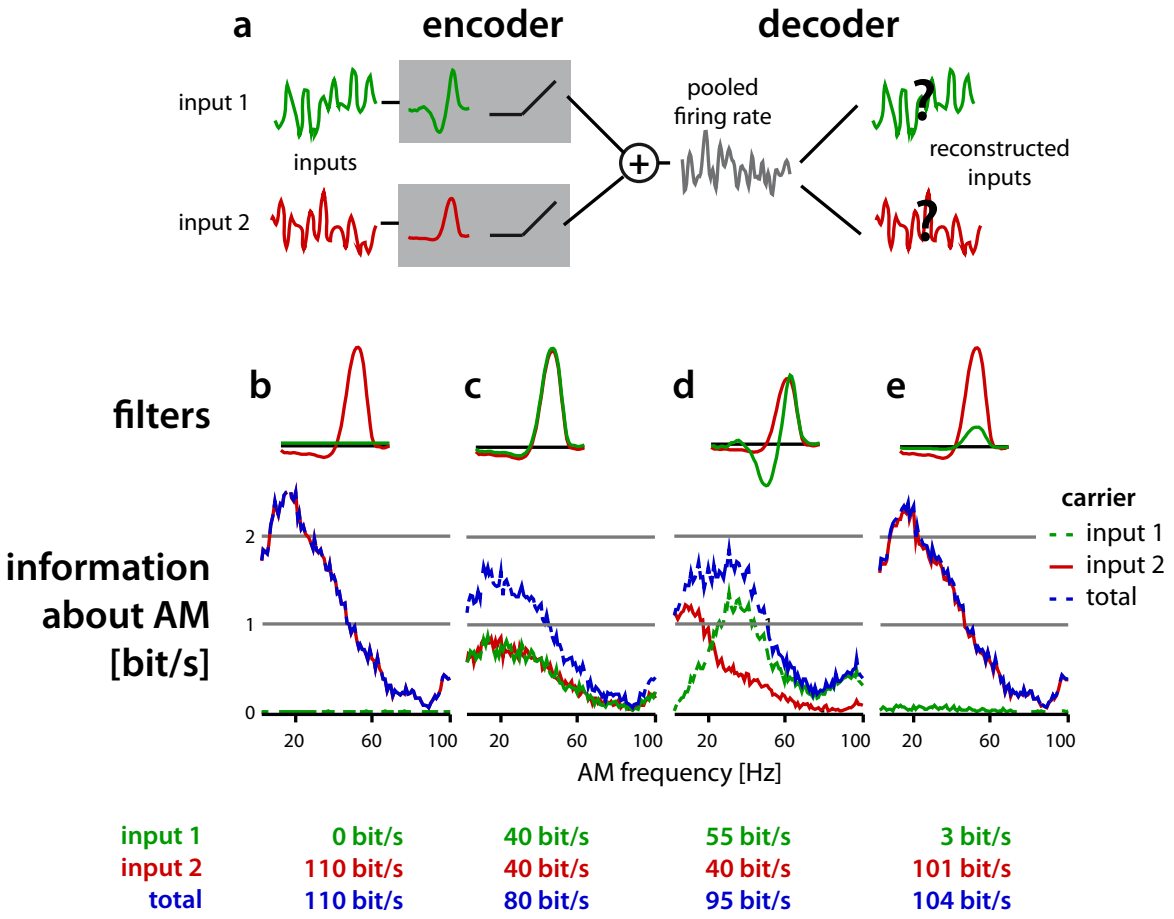

Figure 5.1.: **Abstract model: Effect of suppression and decorrelation of filters on information tuning. a** The encoder models information transmission in a single neuron receiving two, time-varying inputs (green and red lines, respectively). Each input trace is filtered and thresholded independently (grey boxes) before being pooled (grey trace). I then asked how much information an optimal linear decoder can extract from the pooled response about the individual inputs. **b–e** At the top of each panel are the filters for the two input channels. Below that is the information of the neuron as a function of modulation frequency for each input channel (green and red lines) and the total information of the neuron (blue lines, sum of green and red lines). At the bottom is the total information integrated in the interval from 1 to 100 Hz. Four different cases were considered: **b** The neuron receives one input only. **c** The neuron receives two inputs with identical filters for both inputs. **d** Changing one filter to bandpass and **e** suppression of one filter.

Another linear solution of the interference problem was suppression of one of the inputs by reducing the gain of one of the filters (Fig's 5.1 e). This most effectively reduced interference and transmitted maximal information (104 bit/s). Obviously, also this solution entailed a trade-off as only one of the inputs was encoded. This need not be a problem if there are as many encoders as there are relevant inputs: E.g. in the cricket, two frequency bands could be encoded optimally in a two-cell population by complementary suppression.

5.3.2. Implementing the solutions to the interference problem

Both solutions boil down to the intuition that, in order to minimize interference, tuning needs to be narrowed, either for different stimulus components (suppression) or for particular aspects of each stimulus component (decorrelation). While higher selectivity reduces interference for intense and complex stimuli, it also reduces sensitivity, for faint and simple stimuli (Seriès et al., 2004). A sensory system would thus ideally adapt its tuning width to the stimulus statistics, sharpening it for complex stimuli to maximize selectivity and broadening it for simple stimuli to maximize sensitivity. Adaptive tuning and maximization of information transmission has been shown to be implemented through spike-frequency adaptation or gain normalization (Fairhall et al., 2001; Sharpee et al., 2006; Carandini and Heeger, 2011). While the mechanisms underlying adaptive coding are well understood in simple cases, the computations yielding complex stimulus-dependent codes are often elusive (but see Ahrens et al. (2008); Schneider and Woolley (2011)).

How, then, can the two solutions to the interference problem, suppression and decorrelation, be implemented? In a network model inspired by the auditory system of the cricket (Fig. 5.2) I show that both solutions can be obtained using elements ubiquitous in nervous systems: a logarithmic input nonlinearity and broadly-tuned, feed-forward inhibition. This is backed-up by experimental data obtained from the cricket that reproduced the effects observed in the model.

The network model consisted of broadly-tuned high- and low-frequency receptors which encoded the envelope of a sound on a logarithmic decibel scale after linearly weighting different carrier frequencies according to their tuning curve (Fig. 5.2 a–c). These receptors formed the input to a linear three-cell network with connections defined by a linear weight and a temporal delay (Fig. 5.2 d): ascending neuron 1 (AN1) received input only from low-frequency receptors, while ascending neuron 2 (AN2) was most strongly driven by high-frequency receptors. In addition, AN2 also received weak input from low-frequency receptors. The interneuron omega neuron 1 (ON1) got equally strong input from both receptor populations and inhibited the two ascending neurons (Pollack, 1994; Imaizumi and Pollack, 2005; Hirtz and Wiese, 1997).

The encoding properties of different neurons in the modeled network as well as in the recordings were characterized with linear-nonlinear models. This class of model consists of a linear filter which describes a cell's selectivity for temporal features of the envelope. The nonlinearity transforms the filtered stimulus to the firing rate and describes the gain and shape of the cell's tuning for the filter.

First, the system was probed with single carrier frequencies (4.5, 10, 15 or 25 kHz). This

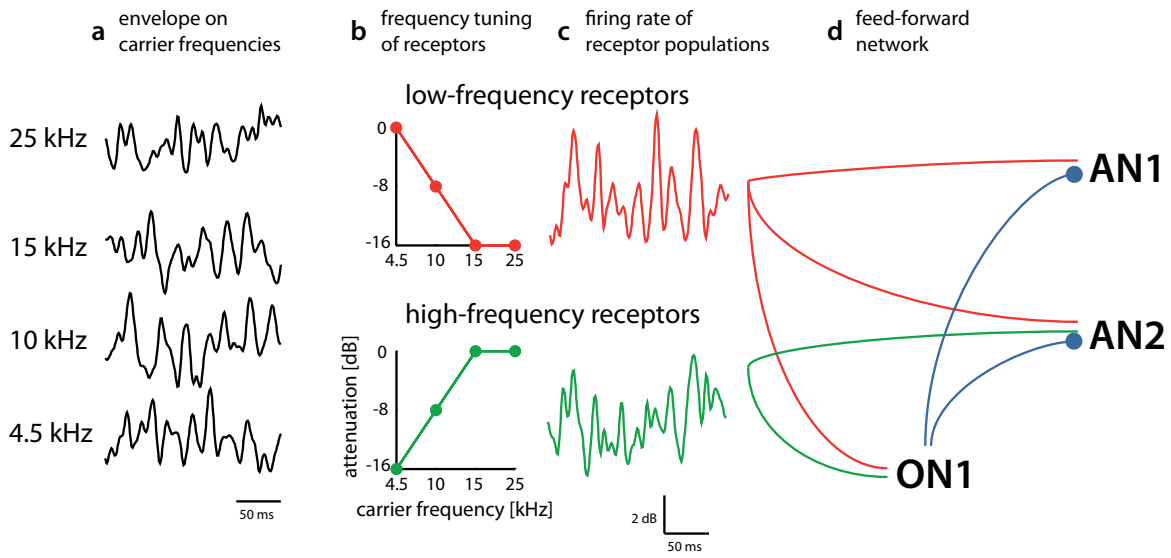

Figure 5.2.: **Network model inpired by the early auditory system of the cricket a** Inputs to the network were defined as the envelope of four carrier frequencies. **b** Two populations of receptors were relatively broadly tuned and were most sensitive to low or high carrier frequencies. Frequency tuning was linear on the sound pressure. **c** The receptors linearily transformed the envelope on the dB to a firing rate. Traces show firing rate resulting from filtering and pooling the envelopes shown in (a). **d** The network integrated inputs from low- and high-frequency receptors (red and green edges, respectively). The ascending neuron AN1 received excitatory input only from low-frequency receptors. AN2 is excited by both receptor populations. ON1 pooled input from both receptor populations and inhibited AN1 and AN2 (blue edges). All neurons were linear rate units. Synapses were implemented with a delay and a linear weight.

yielded one LN model, i.e. a filter and a nonlinearity, for each carrier frequency and cell. Then, the network was stimulated with the four different carrier frequencies simultaneously, each one with an independently modulated envelope. The resulting firing rate was used to estimate a LN model for each individual input, again obtaining one filter-nonlinearity pair for each carrier, but now measured with other carrier frequencies being presented simultaneously. This latter approach was similar to that taken when estimating spectrotemporal receptive fields (STRF's). However, STRF's are usually estimated with a single (or no) nonlinearity for all frequency bands (see e.g. Atencio et al. (2008); Christianson et al. (2008)).

The two solutions to the interference-problem outlined above correspond to changes in the two parts of LN models: for decorrelation, the filters need to change; for suppression, the gain of the nonlinearity needs to change. Interestingly, this system was able to generate both solutions in an adaptive manner, that is, the tuning was broad and temporally uniform for single carrier frequencies and narrow and decorrelated for mixtures of different carrier frequencies.

5.3.3. Adaptive gain

Before coming to changes in the temporal selectivity the changes in the gain of the nonlinearity will be considered. The gain describes how strongly a given input modulates the output firing rate of a neuron. In the linear model neurons, this "nonlinearity" was by design linear; the gain corresponded thus directly to the nonlinearity's slope given by the norm of the filter. In the experimental data, nonlinearities were truly nonlinear and exhibited mostly exponential or sigmoid forms.

For single-carrier frequencies, nonlinearities were relatively uniform (Fig. 5.3 a, b, red lines). In the model they were identical across carrier frequencies for a given cell. In the data, there existed some differences in the shape of the nonlinearities across carrier frequencies, but each cell's firing rate tended to be modulated at all carrier frequencies.

For composite stimuli, nonlinearities differed across carrier frequencies in both AN. These changes were qualitatively similar in the model and the data (compare Fig's 5.3 a and b, black lines). The most drastic dependence of the nonlinearities on stimulus condition appeared in AN1 and AN2, whose firing rates ceased to be modulated by the envelope of some carrier frequencies. This was in stark contrast to the nonlinearities for single-carrier stimuli, which did exhibit a sensitivity to the envelope at all carrier frequencies. In AN1, the nonlinearities for 10, 15 and 25 kHz became almost flat, while that at 4.5 kHz did not change significantly (Fig. 5.3 a, b, middle row). The changes were complementary in AN2, with a strong reduction of the nonlinearity's steepness at 4.5 and 10 kHz, and relatively small changes at the higher carrier frequencies. In the model, I observed an overall reduction of gain at all carrier frequencies in AN2 that was not evident from the data (compare Fig. 5.3 a and b, bottom row, black lines). In the modeled ON1, I also saw an overall reduction of gain, which was—in contrast to both AN—almost uniform across carrier frequencies. In the data, ON1 exhibited a strong reduction of gain at 25 kHz.

To quantify the changes of each cell's selectivity for carrier frequency with stimulus condition, the gain of a cell was quantified as the average of its nonlinearity's slope over all filter values. I then divided each gain by the average gain over all frequencies for that cell and stimulus condition (single or composite). Thereby, relative gain close to 1 indicated little selectivity for carrier frequency (small differences in gain across carrier frequencies); the occurrence of strong deviations from the average gain signaled strong selectivity for carrier frequency. This also corrected for any changes in overall gain.

All three neurons in the network were relatively broadly-tuned for single carrier frequencies (Fig. 5.3 c, d, red). In the model, this gain was identical across frequencies for a given cell (Fig. 5.3 c, red). In the data, gain did change with carrier frequency. However, deviations were relatively small (Fig. 5.3 d, red). Moreover, they were significant only in ON1 and in AN2 at 4.5 kHz. In ON1, this was because the nonlinearity for low but not for high carrier frequencies saturated, yielding a lower average gain despite similar steepness at the beginning of the rising part (Fig. 5.3 b, top row).

In composite stimuli, AN1 and AN2 appeared to be much more selective for carrier frequency than in single-carrier stimuli, both in the model and the data (Fig. 5.3 c, d, black). In contrast, the relative gain of ON1 for composite stimuli remained relatively uniform across frequencies, despite a non-significant reduction of gain at 25 kHz in the data. For AN1, the

relative gain at 4.5 kHz in the data was 2 times greater than the average (almost 3 times in the model). In contrast, the relative gain at 15 and 25 kHz was reduced to 1/2 of the average gain (model: 1/8). In AN2, this effect was complementary to that in AN1. That is, the relative gain in the data at 4.5 kHz was reduced to 1/2 to 1/4 of the average, while that at 15 and 25 kHz was increased to 1.5 times the average gain.

Hence, the tuning of both ascending neurons for carrier frequency became much sharper for composite carrier than for single carrier stimuli. The model was perfectly untuned. The experimental data well approximated this behavior. Tuning arose primarily out of a suppression of the absolute gain at the non-preferred frequencies (as defined by tuning of inputs of the three cells), while the absolute gain at best frequencies did not change much with stimulus condition (Fig. 5.3 b, middle and bottom row). Hence, the ascending neurons responded well to all carrier frequencies when they were presented in isolation. When multiple carrier frequencies were presented simultaneously, they employed a `max`-like and not an average-like encoding of multiple stimuli, as they responded only to their best frequencies and not to the average envelope of all carrier frequencies.

5.3.4. Mechanism for gain changes in the model

I observed adaptive changes in the spectral tuning with stimulus condition in a model of the cricket auditory system as well as in recordings from the same cell types. Note, that the model achieved this adaptive coding without any explicit adaptive mechanisms or changes in its parameters! It is difficult to decisively elucidate the mechanisms underlying this effect *in vivo*. However, in the model one has direct access to the inputs of the three neurons in the model and can thereby locate the source of this adaptive gain. In addition to yielding useful hypotheses that can be tested in future experiments, one will also learn about general mechanisms underlying stimulus-dependent coding.

One candidate mechanism for adaptive coding is the inhibition mediated by ON1, as it gets input from both receptor populations and constitutes thereby a global pool for gain control. However, removing the inhibitory inputs from ON1 to the two AN had only little effect on tuning (Fig. 5.3 c, compare black and green). Evidently, ON1 only sharpened an effect inherited from the receptors. Interestingly, it made tuning more selective despite being broadly-tuned itself. ON1 does this through an iceberg-like effect (Wehr and Zador, 2003; Wu et al., 2008; Isaacson and Scanziani, 2011). Although the strength of inhibition itself does not vary strongly with carrier frequency, that of excitation does through the decrease of excitatory gain at specific frequencies (see below). Hence, the *balance* between excitation and inhibition is altered, resulting in an increase of the relative strength of inhibition at these non-best frequencies.

Having ruled-out ON1 as the origin of stimulus-dependent gain in the model, I conclude that the effect must arise earlier, that is, in the receptors. Indeed, while the output gain of the two receptor populations at different carrier frequencies was as uniform as that of the neurons in the network model, it followed their frequency tuning only for composite stimuli (Fig. 5.4 a).

To understand this unexpected result, I next looked at the only nonlinearity, the transfor-

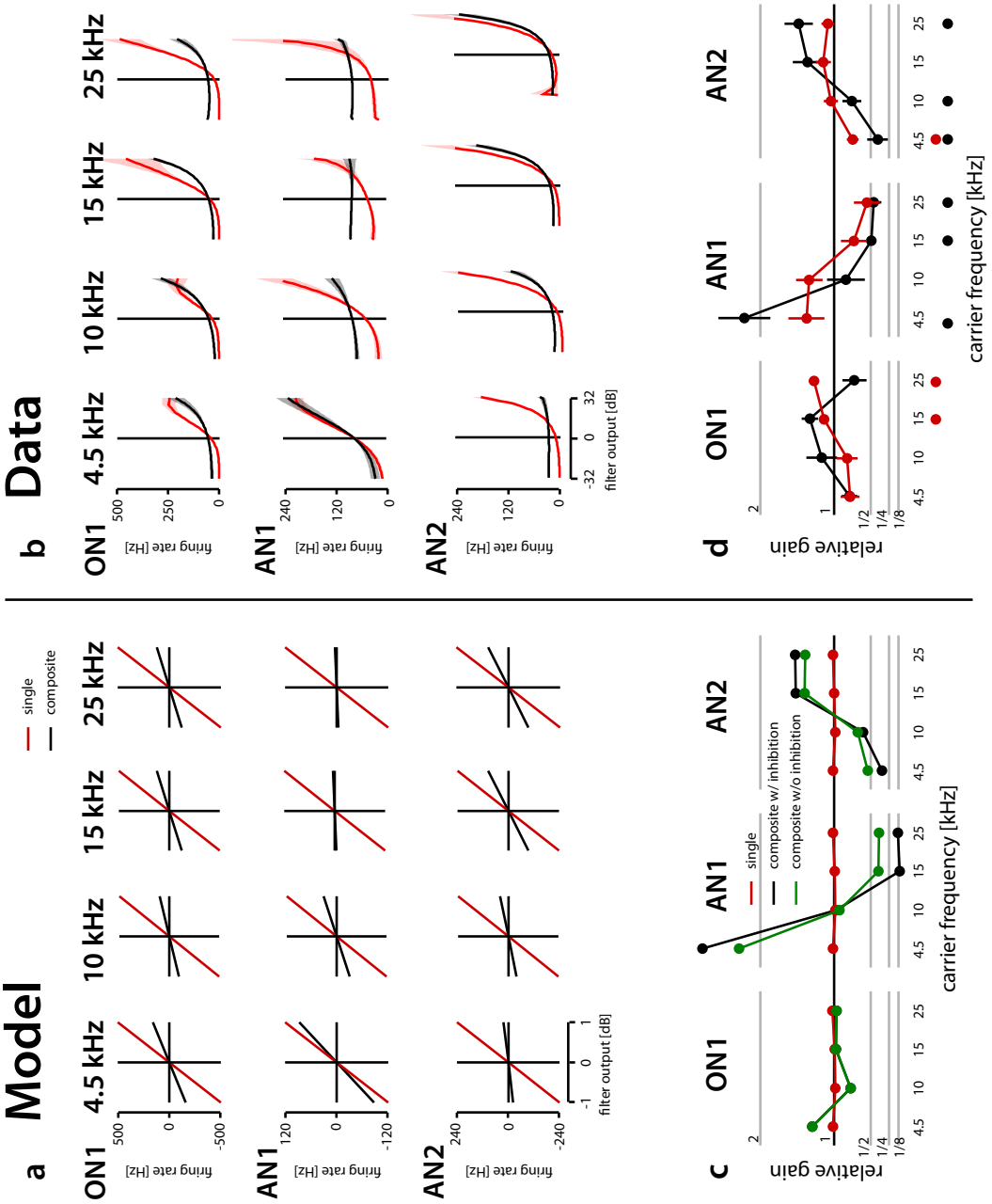

Figure 5.3.: **Model and data: Adaptive changes of gain. a**, **b** The nonlinearities of ON1, AN1 and AN2 in the network model (a) and recorded data (b) for each carrier frequency when measured with isolated ("single", red) or a mixture of four different carrier frequencies ("composite", black). As the model (a) is linear, the "nonlinearities" are straight lines. Experimental nonlinearities exhibit sigmoidal (4.5 kHz) or exponential forms (10, 15 or 25 kHz). Shown is the average $\pm$ s.e.m. (shaded areas) over all recordings. Individual nonlinearities were rescaled by the average firing rate prior to averaging to correct for different response levels. **c**, **d** Gain relative to the average gain across carrier frequencies for ON1, AN1 and AN2 in the model and data for single carrier frequencies (red) and composite carrier frequencies (black). In addition, in the model (c) the relative gain is shown without inhibition from ON1 to AN1 and AN2 (green). Colored dots below the x-axis indicate whether the gain is significantly different from the average (sign rank againt 1.0, $p < 0.05$) for a given frequency in the single (red) or composite (black) condition.

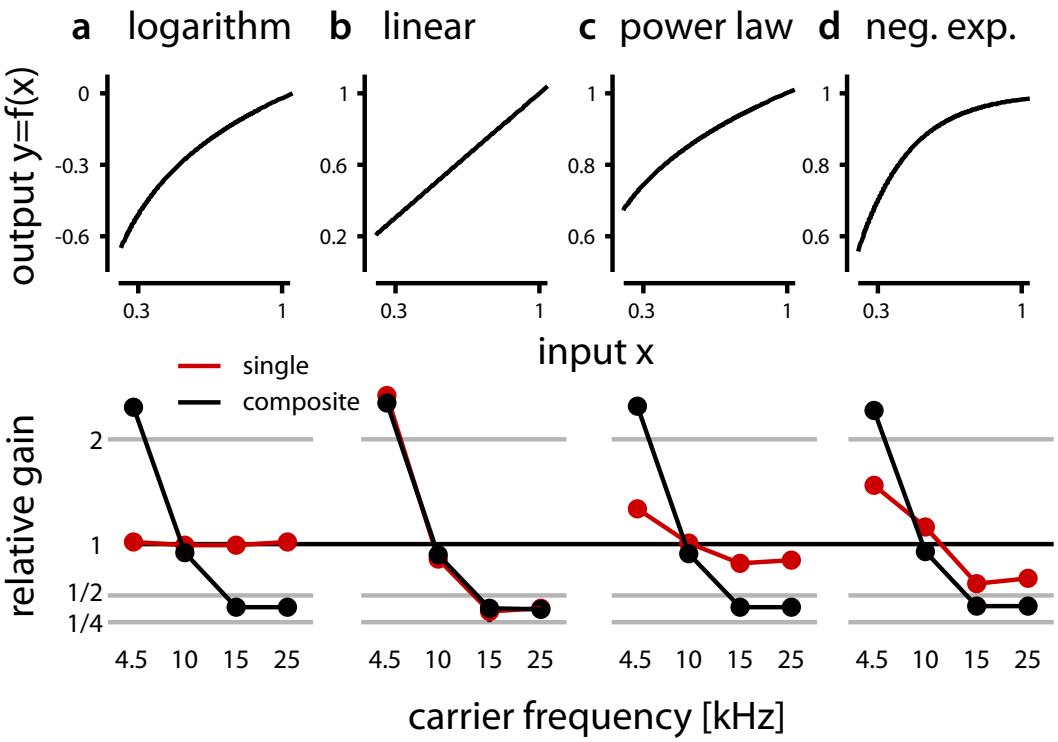

Figure 5.4.: **Model: Many nonlinearities exhibit adaptive gain.** Nonlinearities **a** and the resulting relative gain **b** of low-frequency receptors for single (red) and composite (black) stimuli for different nonlinearities. e1 logarithmic as in the network model - $y = \log(x)$, e2 - linear $y = x$, e3 - power law $y = x^{1/4}$, e4 - negative exponential $y = 1 - \exp(-4x)$.

mation from the pressure to the decibel scale. Replacing this nonlinearity with a linear one, did indeed abolish adaptive tuning: now, all carrier frequencies exhibited differential gain according to the frequency tuning of the receptors in all stimulus conditions (Fig. 5.4 b).

How did the logarithmic nonlinearity enable adaptive tuning? The divisive action of a linear filter on the gain of each carrier frequency had different consequences for single and composite stimuli on a logarithmic scale.

For single-carrier stimuli, the frequency tuning yielded subtractive attenuation on the dB scale. This led to a translation of the stimulus distribution but left the variance of the signal unchanged: $log(p_f \cdot g_f) = I_f + A_f$, where p_f is the pressure at the single carrier frequency f, g_f is the gain of the filter for f, $I_f = log(p_f)$ is the intensity in dB and $A_f = log(g_f)$ is the frequency-specific attenuation. In a static linear system, the output gain corresponds to the input variance. That is, doubling the variance of the input doubles the variance of the output firing rate and hence the output gain. As the frequency-specific tuning of the receptors did not yield changes in the input variance for single carriers (only in their mean), all inputs exhibited equal gain (Fig. 5.4 a, red).

This changes when presenting multiple carrier frequencies. Above equation changes to $log(\sum_f p_f \cdot g_f)$. This expression cannot be factorized into the sum of the log-transformed inputs and a common attenuation factor (if g_f is unequal for different f). In this case, different input components exhibit different weights in the sum according to the frequency tuning g_f of the filter, yielding different, frequency-specific variances in the pooled signal. This differential input gain causes the emergence of differential tuning of excitation, that is a suppression of input gain at strongly attenuated ("non-best") frequencies (Fig. 5.3 a, black).

Note, that the adaptive effect of the nonlinearity did not rely on a specific functional form: a similar effect was observed with power law or exponential-like nonlinearities used to fit the input-output functions of receptors in the olfactory and visual system (Fig. 5.4 c, d). However, only the logarithmic nonlinearity yielded perfectly uniform tuning at single carrier frequencies, as it completely compensated the effects of the multiplicative gain implemented by the frequency tuning of the receptors.

5.3.5. Adaptive changes in filter

Above, we have seen that the balance of inhibition and excitation changed in a stimulus-dependent manner in both the model and the data. To what extent did this affect the temporal selectivity of the neurons in the system? We will see that the relative timing of inhibition and excitation strongly determines the impact of gain changes.

As for the gain, all filters exhibited a similar shape when probed with single carrier frequencies, showing that the neurons in the network were selective for similar temporal features of the envelope at each carrier frequency (Fig. 5.5 a, b, red).

Temporal selectivity of ON1 did not change strongly with stimulus condition (5.5 a, b, top). In contrast, the changes in filter shape for the two ascending neurons were strong and complementary in carrier frequency, parallel to the observations obtained for the gain (compare Fig. 5.3). For AN1, the filters for 4.5 and 10 kHz were unaltered, while filters at 15 and 25 kHz lacked any clear structure (5.5 a, b, middle). This produced large error bars

as the normalization of the filters blew up any random structure in the filters. Some AN1 did lack a well-structured filter even at *single* 15 and 25 kHz, probably because of a sharper tuning or higher threshold of the low-frequency receptors (Fig. 5.6 a)

For AN2, the filters at high carrier frequencies remained unchanged low pass, while the one for the lowest carrier acquired in most cases a bi-modal shape (Fig's 5.5 a, b, bottom). This entails a striking transformation of the coding properties from an encoder of the smoothed envelope at single 4.5 kHz to an encoder of the derivative of the envelope when other carriers were presented simultaneously. This decorrelated the filters for low and high carrier frequencies within AN2. In some recordings of AN2, the filter at 4.5 kHz became fully negatively lobed, probably due to an increase of the strength and/or a decrease of the delay of the inhibition received from ON1 (Fig. 5.6 b).

Thus, the experiments confirmed the expected adaptive and complementary sharpening of tuning for composite stimuli: spectral tuning of AN1 became low-pass with the filter at low carrier frequencies changing only little with stimulus condition. In AN2, the tuning for carrier frequency became sharper and more high-pass; the temporal filter for the envelope at 4.5 kHz changed from a uni- to a bi-modal shape while those for high carrier frequencies did not change.

5.3.6. Mechanism for filter changes in the model

Removing inhibitory inputs from ON1 to both AN in the model abolished the observed changes in filter shape with stimulus condition (Fig. 5.5 c, green). This strong impact of inhibition on filter shape was in contrast to the weak effect of inhibition on gain (Fig. 5.3 c). Interestingly, the gain of ON1 was relatively untuned even for composite carrier stimuli. However, what did decrease strongly for composite stimuli was the excitatory drive to both AN at their respective non-best frequencies. Hence, it was probably not the tuning of inhibition, but that of excitation that changed with stimulus condition and led to changes in filter shape (Fig's 5.3 c, 5.4 a) via a modification in the *balance* of excitation and inhibition. The linearity of the synapses for ON1, AN1 and AN2 in the network model allowed us to decompose the filters of these three neurons into contributions from different inputs and thereby gain a direct insight into how the differential temporal selectivity is created in the model.

The relative strength of inhibition did indeed increase for non-preferred, but not for preferred carrier frequencies (Fig. 5.7, blue lines). E.g. when comparing the relative strength of inhibition of AN1 between single and composite stimuli, we see that it increased for 15 kHz (Fig's 5.7 b)—the cell's "non-preferred" carrier frequency—but not for 4.5 kHz—its "preferred" carrier frequency (Fig's 5.7 a). Hence, the changed balance between excitation and inhibition led to a stronger impact of inhibition on filter shape at non-preferred carrier frequencies in the composite condition.

However, it was the timing of inhibition relative to the excitation that determined its specific effect. Since excitation and inhibition were coincident in AN1 (Pollack, 1994), they canceled each other out if balanced and produced thereby an unstructured filter at non-preferred frequencies (e.g. 15 kHz, Fig's 5.5 b, black line). This ensured that AN1 did not

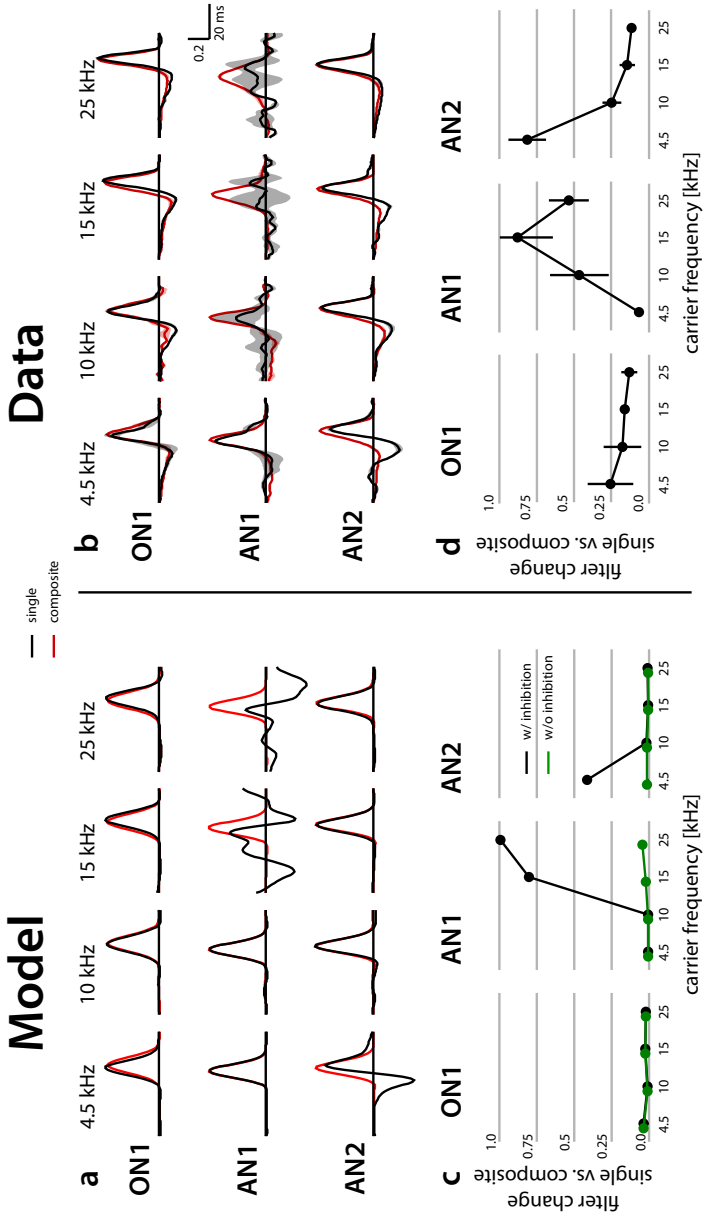

Figure 5.5.: **Model and data: Adaptive changes in filter shape. a**, **b** Filters for ON1, AN1 and AN2 (rows) for different carrier frequencies (columns) when presented in isolation ("single", red) or as a sum ("composite", black). Filters were normalized to unit norm. Experimental filters (b) represent average ± s.e.m. (shaded areas) over individual recordings. Experimental filters were aligned at their peaks for better visibility. Note that the s.e.m. is usually smaller than the line thickness for most filters except for AN1 at 10, 15 and 25 kHz as the individual filters lacked any clear structure. **c**, **d** Changes in filter shape when comparing single- and composite-carrier filters for ON1, AN1 and AN2 in the model (c) and the data (d). Change was calculated as one minus the dot product of the normalized filters for each condition and carrier frequency. Dots and error bars in (d) represent average ± s.e.m. of the change across individual recordings. Green plots in (c) show the changes in filter shape in a model without inhibition from ON1 to AN1 and AN2.

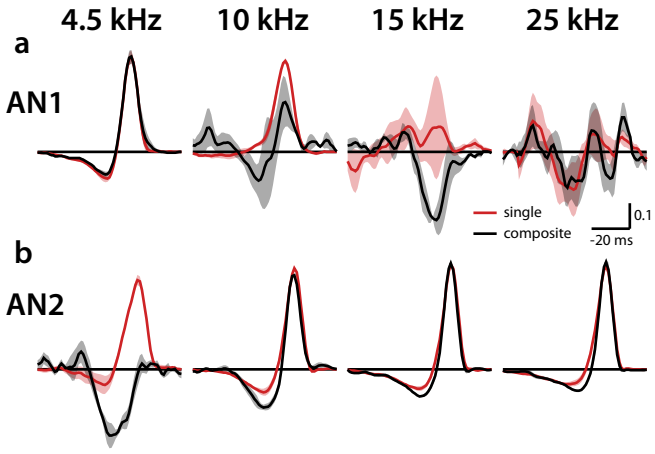

Figure 5.6.: **Data: Special variants of AN1 and AN2. a** While many specimen of AN1 exhibited considerable gain and well-structured filters at high carrier frequencies, 3/6 of the individual cells did not lock well to features of the envelope at 15 and 25 kHz even if they were presented in isolation. This was probably due to strong attenuation of low-frequency receptors at high carrier frequencies or due to strong inhibitory inputs from ON1 (see negative filter at 15 kHz). **b** For AN2, most filters exhibited a bi-modal shape at 4.5 kHz for composite-carrier stimuli. For 4/12 few cells, we observed a fully negative filter, indicating that stimulus power at 4.5 kHz suppressed firing probably due to a stronger inhibitory input from ON1.

respond to high carrier frequencies in a composite stimulus, even if their amplitude was high.

In AN2, inhibition was slower than high-frequency excitation (Marsat and Pollack, 2007). This *non*-coincidence of the different inputs resulted in a bimodal filter: the high-frequency excitation was faster then both the low-frequency excitation and the inhibition (Pollack, 1994). Thus, excitation initially outweighed inhibition, leading to the positive lobe of the filter. Eventually, a negative lobe emanated due to the slow inhibition out-weighing low-frequency excitation (Fig's 5.7 c).

5.3.7. Stimulus-dependent coding preserves information in the cricket

Our experimental data show that the auditory system of the cricket exhibits stimulus-dependent tuning (Fig's 5.5 b and 5.3 b). I quantified information in the experimental data to confirm the predictions of the abstract encoder model (see Fig. 5.1). That is, information should be preserved if all but one inputs are suppressed or if the filters for all inputs are decorrelated.

Information at 4.5 kHz in AN1 was unaffected by the presence of multiple carrier frequencies (Fig. 5.8, middle, $p = 0.13$, compare Fig. 5.1 e). This confirms that the suppression of output gain for higher carrier frequencies can reduce interference: While AN1 transmitted some, though little, information about the envelope at 10, 15 or 25 kHz when they were presented in isolation, almost no information was transmitted about these carrier frequencies in a broadband stimulus (Fig. 5.8, middle, compare red and black lines).

In the other two neurons in the network, AN2 and ON1, the loss of information was substantial at all carrier frequencies as suppression strongly affected only one or two carrier

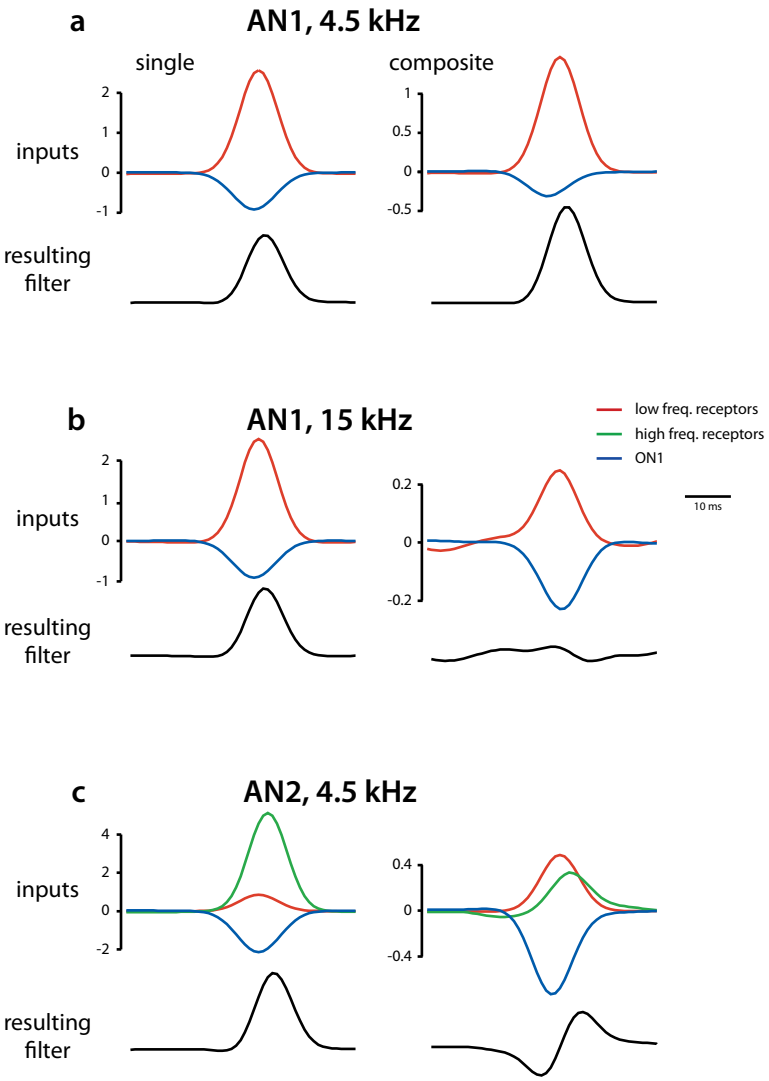

Figure 5.7.: **Model: Balance between inhibition and excitation and the timing of inhibition shapes adaptive change in filter shape.** "Generation" of the filter shapes (black) in AN1 (a, b) and AN2 (c). Shown are the filter for the inputs, multiplied by the synaptic weight and delayed by the synaptic delay (low frequency receptors - red, high frequency receptors - green, ON1 - blue) and their changes with stimulus condition (left - single, right - composite). The resulting filters are the sum of the the weighted and delayed input filters and have the same scale as the inputs shown above. **a** AN1 at 4.5 kHz (preferred carrier frequency), **b** AN1 at 15 kHz (non-preferred), **c** AN2 at 4.5 kHz (non-preferred).

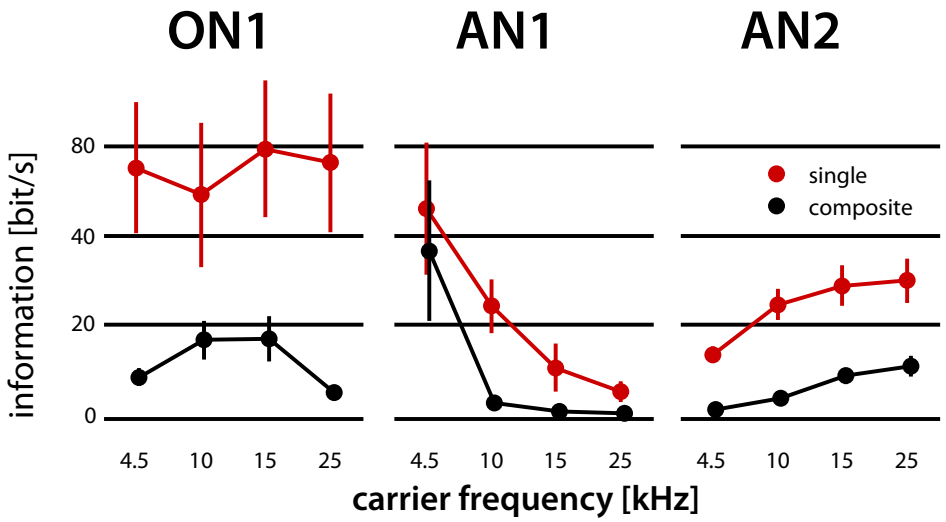

Figure 5.8.: **Data: Information about the envelope of different carrier frequencies.** Shown is the coherence information in recorded responses about the envelope at different carrier frequencies when presented in isolation (single - red) and as a mixture (composite - black). Plots show mean ± s.e.m.

frequencies (Fig. 5.8, left and right). Thus, many carriers were left to interfere with each other, leading to a large overall loss of ≈ 70% of the information about individual stimulus components in these two cells when comparing single and composite-carrier stimuli (compare Fig. 5.1 c).

Interestingly, the decorrelation of filters in AN2 had little effect on information transmission in the data (Fig. 5.8 right, compare Fig. 5.1 d), probably because the gain at 4.5 kHz was strongly suppressed (Fig. 5.3 b). This was mostly because suppression of gain at this carrier frequency overrode the decorrelation of filters. Hence, suppression had the dominant effect on information transmission in the data. Although suppression leads to a loss of information about all but one input component, this solution is efficient in a system with two relevant frequency bands and two output channels as instantiated by the auditory system of the cricket.

5.4. Discussion

I have shown that a network with a logarithmic input nonlinearity and feed-forward inhibition as implemented in the auditory system of the cricket adaptively sharpens its tuning when confronted with multi-component stimuli.

Information about individual components in a mixture can be preserved by decorrelating the tuning for carrier frequency and/or temporal features at each carrier frequency: For single-carrier stimuli, both the gain as well as the filters are relatively uniform across carriers, leading to correlated firing of both ascending neurons. In contrast, for composite-carrier stimuli both the gain and the filters change in a complementary manner, yielding decorrelated firing of both ascending neurons to a broadband stimuli (Fig's 5.5 b and 5.3 b). Thus, in addition to reducing redundancy, decorrelated coding may also benefit coding by reducing the interference between the representation of different stimulus components (compare Barlow (2001); Vinje and Gallant (2000); Wiechert et al. (2010)).

Stimulus-dependent tuning has been shown for many systems (narrower tuning for global vs. local stimuli in electric fish, Chacron et al. (2003); spectrally dense stimuli in A1, Blake and Merzenich (2002); stimuli with strong spatial correlations in the visual system, Vinje and Gallant (2000)) and is usually taken as an indicator for nonlinear coding (Machens et al., 2004). Nonlinear, stimulus-dependent coding is often attributed to subthreshold or inhibitory inputs and the spiking threshold (Schneider and Woolley, 2011; Priebe and Ferster, 2008). These mechanisms change the balance of excitation and inhibition in a stimulus-dependent manner. Our results suggest that the same effects can be achieved by a peripheral mechanism without any spiking threshold: the logarithmic transformation of sound amplitude to the dB scale by receptors. This yields stimulus dependent tuning of excitation—broad for single-carrier and narrow for composite-carrier stimuli (Fig. 5.4 a)—and thereby yields the observed changes in output gain and filter shape (Fig's 5.5 and 5.3).

A similar phenomenon would also be observed in standard, filter bank models of the cochlear as long as different stimulus components fall into the same "critical band" (see e.g. Zilany et al. (2009); Fontaine et al. (2011)). This could partially account for the stimulus or context dependence of spectrotemporal receptive fields in primary auditory cortex (Gourévitch et al., 2009; Pienkowski and Eggermont, 2011): tuning for single-tone or narrow-band stimuli is much broader than that of wide-band stimuli. Our simple model suggests that the logarithmic peripheral nonlinearity in the cochlea might at least partially contribute to this in addition to synaptic depression or broadly-tuned inhibition usually proposed (Eggermont, 2010).

Moreover, as this mechanism does not depend on a specific functional form of the input nonlinearity (see Fig. 5.4), it could potentially be at work in any encoding system with a compressive nonlinearity after a linear filter stage (e.g. concentration in the olfactory system Olsen et al. (2010); brightness and contrast in the visual system, e.g. Sharpee et al. (2006)). It could even by implemented in a single neuron integrating different inputs from weighted synapses in a saturating dendritic tree (Prescott and De Koninck, 2003).

In addition to the change in the balance of excitation and inhibition through the logarithmic nonlinearity, broadly-tuned inhibition further sharpened tuning and change the tem-

poral selectivity of cells in a stimulus-dependent manner (Fig. 5.3 c). In the model, this is mediated by ON1, which pools activity of all receptors and represents thereby the full input to AN1 and AN2. ON1 is usually associated with the enhancement of directional cues in the system (Schildberger, Marsat and Pollack (2005, 2007)) but has also been shown to shape temporal selectivity in the cricket (Tunstall and Pollack, 2005). This study interprets its action as mediating a global gain control signal to the ascending neurons. While this is conveyed by a single neuron, ON1, in the cricket auditory system, such signals are usually assumed to be implemented in the form of pools of many neurons (Rothman et al., 2009; Gabbiani et al., 2002). However, recently, a *single*, giant inhibitory interneuron in the mushroom bodies of locusts and Drosophila has been reported to represent the global gain of the population of Kenyon cells (Papadopoulou et al., 2011). Global gain control also is known to be involved in increasing sparseness, controlling a network's overall activity or conveying selectivity for complex stimuli (Carandini and Heeger, 2011; Vinje and Gallant, 2000; Silbering and Galizia, 2007). Our results with a simple model have shown that untuned inhibition contributes to the decorrelation and sharpening of tuning through an iceberg-like effect just as has been reported in cortex (Wehr and Zador, 2003; Wu et al., 2008; Isaacson and Scanziani, 2011; Wilent and Contreras, 2005; Luo et al., 2010; Xing et al., 2011). Interestingly, the effect of inhibition from ON1 on AN2 is similar to that observed in barrel cortex: it is coincident with non-preferred stimuli and thereby more effective in suppressing these (Fig. 5.7 c, see Wilent and Contreras (2005)).

While this simple model was successful in reproducing suppression and decorrelation as observed in the experimental data, biological networks—including the cricket auditory system—are much more complex and are equipped with a wealth of nonlinearities. How the above mechanisms interact with a, possibly dynamic, spiking threshold, spike-frequency adaptation, presynaptic inhibition or other sources of inhibition (Benda and Hennig, 2008; Wimmer et al., 2008; Pollack, 1988; Hildebrandt et al., 2011; Nolen and Hoy, 1986b) is not entirely clear. E.g. fast inhibition to AN2 that could not be identified with ON1 has been reported to underly two-tone inhibition in *Teleogryllus oceanicus*. However, there is no evidence for fast inhibition in our experiments with the distantly related cricket species *Gryllus bimaculatus*.

Another—nonlinear—solution to the problem of encoding multiple stimuli in the cricket has been reported previously: AN2 produces bursts of action potentials specifically in response to ultrasound but not loud, low-frequency sounds (Marsat and Pollack, 2006). Thereby, high carrier frequencies are encoded via short interspike intervals or high frequencies in the response and low-carrier frequencies are encoded via low response frequencies. This burst code is nonlinear, as the intra-burst frequency does not necessarily equal the stimulus frequency by which the burst is evoked (Theunissen and Miller, 1995).

The coding found in the cricket is reminiscent of max-like as opposed to average-like encoding of multiple stimulus components: that is, both AN1 and AN2 respond to the strongest inputs in a mixture, not to the average (compare Gawne and Martin (2002); Lampl et al. (2004); Carandini and Heeger (2011) and van Wezel et al. (1996); Zoccolan et al. (2005)). This seems efficient, as a motion-sensitive neuron in area MT should not respond to the average motion of two divergent objects in the visual field; likewise, AN1 or AN2 should

not respond the average envelope across all carrier frequencies, because different carrier frequencies convey different meaning to the cricket.

There are sensory systems, for which the problem of encoding multiple stimuli poses different challenges: In olfactory systems or the song bird auditory system, stimuli are identified on the base of a complex mixture of different stimulus components. In this case, adaptive sharpening of tuning in early stages of sensory processing might enable a faithful encoding of individual components, combinations of which can then be detected by `and`-like coincidence detection (Margoliash and Konishi, 1985; Riffell et al., 2009; Borst et al., 2005).

6. Conclusion

In this thesis, I studied sparse and adaptive coding in insect auditory systems and covered a large range of phenomena including

1. the transformation of the population code in grasshoppers (chapter 2),
2. the contribution of nonlinear computations to temporal and population sparseness (chapter 3),
3. a simple mode of song recognition (chapter 4),
4. stimulus-dependent coding and the role of a logarithmic nonlinearity and inhibition (chapter 5).

The implications of the results for sparse and adaptive coding have been discussed in each of the chapters. I will now briefly recapitulate the main results of the thesis, highlight some common themes that reoccurred throughout and provide a general outlook.

6.1. Sparse coding

Part I has dealt with sparse coding. I established to what degree sparseness arises in the auditory system of grasshoppers (chapter 2), revealed the abstract computations leading to sparseness (chapter 3), and proposed and tested a hypothesis about the functional significance of this transformation (chapter 4).

The early auditory system of grasshoppers creates a temporally and population sparse representation of song after two processing stages from dense and uniform inputs. While in the first two processing stages (receptors and local neurons), all neurons respond to very similar features, neurons in the output stage are selective for very specific features. Using a population decoder, I could show that this leads to a transformation of the population code from a summed-populations to a labeled-line code. That is, a spike's neuron of origin—its label—becomes informative at the output stage of the network. This yields an explicit representation of temporal features of the song, where the presence or absence of spikes in a particular ascending neurons signals the presence or absence of a specific temporal feature in the stimulus. Although the degree of sparseness in the auditory system of the grasshopper is smaller than that of ultra-sparse codes as found in e.g. the Kenyon cells of the olfactory system of locusts (Perez-Orive et al., 2002), it is still surprising to find such a complex transformation of the population code at an early stage of processing in such a small system.

Applying spike-triggered covariance analysis to recordings from second and third-order neurons, I found two non-linear computations contributing to the temporal and population sparseness measured in the system: A sensitivity to the derivative of the stimulus and the `and`-like combination of an excitatory and a delayed suppressive stimulus feature. Both accentuated responses to stimulus transients and increased temporal sparseness. The latter generated a large diversity of temporal filters through a wide range of delays between the excitatory and suppressive feature and thereby increased population sparseness. Incorporating prior knowledge about the neurons in my data set, I proposed putative biophysical mechanisms underlying both computations: adaptive currents and leading inhibition, respectively. These are mechanisms found in many neural systems—the computations proposed here are therefore likely to contribute to sparseness in other organisms as well (see e.g. Farkhooi et al. (2009)). Blocking inhibitory synaptic transmission or adaptation currents should decrease the temporal and population sparseness of the system (compare Römer and Seikowski (1985)).

In addition to an increase of population sparseness, I also found that the temporal fidelity of the neural representation degraded in the output stage of the network (see also Vogel et al. (2005)). The system seemed thus to make a trade-off between information about the nature and the timing of features. In conjunction with evidence from previous behavioral experiments (von Helversen and von Helversen, 1998) this led me to hypothesize that neurons in the grasshopper's brain that receive input from the ascending neurons do not evaluate the timing of spikes but rely primarily on spike count. As the labeled-line population code at that stage entails an explicit representation of the presence or absence—or the scalar value—of temporal features of a song, this might be enough information to recognize

the correct species and quality of the sender. This would greatly facilitate song recognition, as it would only entail integration of the output of the metathoracic network over time and no extraction of information based on spike timing. That way, the population-sparse labeled-line code at the output of the network would enable a very simple algorithm for the evaluation of the long communication signals in grasshoppers.

I designed a classifier, inspired by the properties of the auditory system of grasshoppers that ignored the timing of features and based its decision only on the time-average output of feature detectors. This classifier was trained on behavioral data using a genetic algorithm and I could show that most of the behavioral responses could be explained by such a simple mode of song recognition. This indicates that a complex task—the evaluation of a long communication signals—can indeed be efficiently solved using this simple algorithm. However, a *single* temporal feature was sufficient to predict behavior well, which is somewhat at odds with the diversity of feature selectivity found at the level of ascending neurons. Although this does not directly contradict the notion that population sparseness enables this simple mode of song recognition—it could well be a limitation of the data sets— it does not support it either. Hence, the hypothesis that population sparseness facilitates song recognition in the grasshopper needs more data to be confirmed or rejected. In addition, the classifier performed relatively poorly on a data set consisting of conspecific natural signals, primarily due to the limited variability of the stimuli and response values. Extending this data set with heterospecific songs will increase both the diversity of stimulus patterns as well as the range of response values and will show that the classifier performs well for natural signals as well.

6.2. Stimulus-dependent coding

Part II dealt with adaptive, stimulus-dependent coding in the cricket. I have found that the two ascending neurons are tuned relatively broadly when stimulated with single-carrier stimuli. That is, both their temporal selectivity as well as their selectivity for carrier frequency is relatively uniform across carrier frequencies. However, tuning changes drastically when a mixture of four carrier frequencies is presented: while the frequency tuning sharpens, the temporal tuning becomes more specific to the carrier frequency. This occurs in a complementary way, that is, AN1 suppresses responses to high carrier frequencies and AN2 suppresses responses to low carrier frequencies. That a simple system is able to display such complex, nonlinear phenomena is surprising.

Using an abstract encoder model, I showed that this mode of stimulus-dependent coding is able to preserve information about individual stimulus components: A broadly tuned encoder will respond to multiple stimulus components (i.e. carrier frequencies). This leads to interference and ambiguity, and consequently to a loss of information about individual stimulus components. Narrowing the tuning reduces interference, as an ideal encoder responds now to a single stimulus component only. In fact, the two solutions possible in a linear system (suppression and decorrelation) are implemented in the auditory system of the cricket although only one of them (suppression) is effective in preserving information.

Furthermore, a simplified network of the cricket's auditory system revealed two general

mechanisms underlying stimulus-dependent coding: a logarithmic nonlinearity implied by a representation of sound amplitude on a dB scale and broadly-tuned inhibition. The logarithmic nonlinearity in the model is implemented after a linear tuning function for carrier frequency in the receptors and enables adaptive tuning of excitatory inputs to the network, that is broad tuning for single and narrow tuning for multi-carrier stimuli. Inhibition is broadly-tuned irrespective of stimulus condition. However, the adaptive tuning of excitatory inputs changes the balance of inhibition and excitation in a stimulus-dependent manner, increasing the impact of inhibition for particular frequencies and thereby decorrelating the temporal selectivity of both AN for multi-carrier stimuli. The predictions of this model are two-fold: First, adaptive spectral tuning is inherited by the receptors. This could be tested by recording the population activity of auditory receptors. Second, inhibition mediated by ON1 is necessary to adaptively change the temporal selectivity. Blocking inhibitory synaptic transmission should therefore abolish these changes (see e.g. Skiebe et al. (1990)). As the adaptive effect of the logarithmic nonlinearity does not depend on a specific functional form and as broadly-tuned inhibition is found in many neural system, the proposed mechanisms could enable stimulus-dependent coding in other systems as well.

6.3. Concluding remarks

I now want to highlight some recurrent themes and connections between the different chapters and end with a brief outlook.

Specificity and labeled lines The decorrelation of feature selectivity in the grasshopper (chapter 2) changes the "meaning" of synchronous firing. In local neurons, population synchrony reflects the co-tuning of cells in early stages of the network. Synchrony at the level of ascending neurons signals the co-occurrence of different stimulus features. Decorrelation leads to population sparseness and a transformation of the population code to a "labeled line", as each output neuron encodes the presence or absence of a specific temporal feature of the stimulus to read-out neurons. Interestingly, a similar theme appeared when asking about the optimal "population" code in the auditory system of the cricket, when it is confronted with multi-component stimuli (chapter 5). In an abstract model of the network, I have found a stimulus-dependent or "conditional" labeled-line, as the two ascending neurons of the cricket adaptively sharpened their spectral tuning in a complementary manner.

Inhibition and decorrelation Modeling in both the grasshopper and the cricket suggests that inhibition contributes to generating this labeled line code (chapters 3 and 5). In the grasshopper, the large range of delays between "excitatory" and suppressive stimulus features produced a high diversity of temporal filters, which were essentially uncorrelated. In the cricket, the model suggests that stimulus-dependent changes in the balance between excitation and inhibition decorrelate the temporal and—to a lesser extent—the spectral selectivity of the two output neurons in a complementary manner, again leading to a decorrelation of responses.

Moreover, the song-recognition circuit has shown that suppressive features, which could be implemented using inhibition, also increase the selectivity for song (chapter 4): While the individual features were rather broadly-tuned (low-pass and weak band-pass), linear combination of the two filters yielded a strong band-pass filter. Most interestingly. The output of the individual filters exhibited only little correlation with behavior. This has important consequences when searching for the neuronal correlate for behavior: Usually, one seeks neurons whose response characteristics correlate well with behavioral tuning. The results of the classifier suggest that in a population code, this need not necessarily be the case and that the combined response of all neurons in a population should be taken into account.

The powers and limitations of linearity Throughout this thesis, we have seen how far one gets with linear descriptions of neural computations. The appeal of linear models is the that they are relatively simple to fit, require little data for their estimation and are easy to interpret.

I have successfully employed linear-nonlinear models which are linear in the filter to fit auditory neurons in the grasshopper and cricket. Moreover, a linear network model successfully reproduced stimulus-dependent coding in the cricket.

Along this line, the classifier for explaining song recognition in grasshoppers relied on linear-nonlinear cascades for feature detection. Additionally, different feature values were weighted linearly to predict the behavioral response. This linear approach well reproduced the complex—and non-monotonous—interactions between different features of a song. However, in all these cases, an additional, nonlinear part was necessary to fully account for experimental findings: be it a nonlinearity following the linear filter in LN models or a logarithmic input nonlinearity for the network model.

One of the main tasks of nonlinearities was to convey high selectivity when integrating multiple stimulus features (chapters 3 and 5). In the ascending neurons of grasshoppers, this was the `and`-like integration of excitatory and suppressive features, making the neuron selective for a small subset of feature combinations. This increased temporal and population sparseness. In the cricket, the logarithmic input nonlinearity adaptively sharpened the tuning of the inputs to a linear network model, thereby increasing specificity in the presence of multi-component stimuli.

Although a linear combination can also increase selectivity, as we have seen in the combination of a low-pass and a weak bandpass filter yielding a strongly tuned bandpass filter (chapter 4, Fig. 4.8), nonlinearities entail thresholds, saturation, adaptation. This enables higher selectivity by allowing much sharper and flexible boundaries between stimulus categories (spiking vs. non-spiking, response vs. no response).

Conclusion and outlook In this thesis, I have shown that insect auditory systems are simple, but not simplistic. That, despite their small size and restricted tasks, they can exhibit sophisticated processing capacities. Thanks to their small size and restricted tasks, they can provide valuable insights into how neural computation works and inform hypotheses for equivalent computations in more complex nervous systems.

Among those are the insight that song recognition can benefit from ignoring an aspect of information—the temporal structure of a stimulus. This principle seems to be at work in the grasshopper. It could be combined with the classifier approach to explore perceptual decision making in other animals and sensory modalities.

Also, the study of adaptive coding has revealed a surprisingly simple computation—a logarithmic nonlinearity—which can provide stimulus-dependent gain for different stimulus components. Other sensory systems could by tested for the presence of the same phenomenon.

By looking at computations in the periphery, I have developed and tested hypotheses about how the brain could process inputs from earlier processing stages. The representation of song in the brain likely relies on the spike count of ascending neurons (see e.g. Römer and Seikowski (1985)). However, the features represented thereby need not necessarily be correlated with behavior. These insights could inform further experimental studies in the grasshopper brain.

Using computational methods to analyze data, I described the transformation of the neural codes along a neural signal transduction chain or in dependence of stimulus conditions. Together with abstract models, this has revealed computations underlying these transformations. While this has proven a powerful approach, it provided only limited insight into the specific biophysical mechanisms implementing these computations.

Although prior knowledge about the systems under examination helped constrain hypotheses about mechanisms, only further experiments in combination with more detailed and realistic modeling can advance our understanding. E.g. pharmacology to switch-off specific types of synapses or currents might help to determine whether the computations underlying sparseness are cell-intrinsic or network mechanisms. Or, by extending the model of the auditory system of the cricket from a phenomenological/algorithmic level to a more biophysically realistic, spike-based one might reveal details about the implementation of stimulus-dependent coding.

List of Figures

List of Tables

Bibliography

Ahrens, M. B., Linden, J. F., and Sahani, M. (2008). Nonlinearities and Contextual Influences in Auditory Cortical Responses Modeled with Multilinear Spectrotemporal Methods. *Journal of Neuroscience*, 28(8):1929–1942.

Ahumada, A. and Lovell, J. (1971). Stimulus Features in Signal Detection. *The Journal of the Acoustical Society of America*, 49(6B):1751–1756.

Amézquita, A., Flechas, S. V., Lima, A. P., Gasser, H., and Hödl, W. (2011). Acoustic interference and recognition space within a complex assemblage of dendrobatid frogs. *Proceedings of the National Academy of Sciences of the United States of America*.

Aronov, D., Reich, D. S., Mechler, F., and Victor, J. D. (2003). Neural coding of spatial phase in V1 of the macaque monkey. *Journal of Neurophysiology*, 89(6):3304–3327.

Assisi, C., Stopfer, M., Laurent, G., and Bazhenov, M. (2007). Adaptive regulation of sparseness by feedforward inhibition. *Nature Neuroscience*, 10(9):1176–1184.

Atencio, C. A., Sharpee, T. O., and Schreiner, C. E. (2008). Cooperative Nonlinearities in Auditory Cortical Neurons. *Neuron*, 58(6):956–966.

Balakrishnan, R., von Helversen, D., and von Helversen, O. (2001). Song pattern recognition in the grasshopper Chorthippus biguttulus: the mechanism of syllable onset and offset detection. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 187(4):255–264.

Barlow, H. (2001). Redundancy reduction revisited. *Network: Computation in Neural Systems*, 12(3):241–253.

Benda, J., Bethge, M., Hennig, R. M., Pawelzik, K. R., and Herz, A. V. M. (2001). Spike-Frequency Adaptation: Phenomenological Model and Experimental Tests. *Neurocomputing*, 38-40:105–110.

Benda, J. and Hennig, R. M. (2008). Spike-frequency adaptation generates intensity invariance in a primary auditory interneuron. *Journal of Computational Neuroscience*, 24(2):113–136.

Blake, D. T. and Merzenich, M. M. (2002). Changes of AI receptive fields with sound density. *Journal of Neurophysiology*, 88(6):3409–20.

Borst, A., Flanagin, V. L., and Sompolinsky, H. (2005). Adaptation without parameter change: Dynamic gain control in motion detection. *Proceedings of the National Academy of Sciences of the United States of America*, 102(17):6172–6.

Borst, A. and Theunissen, F. E. (1999). Information theory and neural coding. *Nature Neuroscience*, 2(11):947–957.

Boyan, G. S. (1999). Presynaptic contributions to response shape in an auditory neuron of the grasshopper. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, V184(3):279–294.

Brenner, N., Bialek, W., and Ruyter, D. (2000). Adaptive rescaling maximizes information transmission. *Neuron*, 26(3):695–702.

Brette, R. (2010). On the interpretation of sensitivity analyses of neural responses. *The Journal of the Acoustical Society of America*, 128(5):2965.

Brody, C. D. (1999). Disambiguating different covariation types. *Neural Computation*, 11(7):1527–1535.

Bugrov, A., Novikova, O., Mayorov, V., Adkison, L., and Blinov, A. (2005). Molecular phylogeny of Palaearctic genera of Gomphocerinae grasshoppers (Orthoptera, Acrididae). *Systematic Entomology*, 31(2):362–368.

Carandini, M. and Heeger, D. J. (2011). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62.

Chacron, M. J., Doiron, B., Maler, L., Longtin, A., and Bastian, J. (2003). Non-classical receptive field mediates switch in a sensory neuron's frequency tuning. *Nature*, 423(6935):77–81.

Chacron, M. J., Longtin, A., and Maler, L. (2011). Efficient computation via sparse coding in electrosensory neural networks. *Current Opinion in Neurobiology*, 21(5):752–60.

Chacron, M. J., Maler, L., and Bastian, J. (2005). Electroreceptor neuron dynamics shape information transmission. *Nature Neuroscience*, 8(5):673–8.

Chittka, L. and Niven, J. (2009). Are bigger brains better? *Current Biology*, 19(21):R995–R1008.

Christianson, B. G., Sahani, M., and Linden, J. F. (2008). The Consequences of Response Nonlinearities for Interpretation of Spectrotemporal Receptive Fields. *Journal of Neuroscience*, 28(2):446–455.

Clemens, J., Kutzki, O., Ronacher, B., Schreiber, S., and Wohlgemuth, S. (2011). Efficient transformation of an auditory population code in a small sensory system. *Proceedings of the National Academy of Sciences*, 108(33):13812–13817.

Clemens, J., Weschke, G., Vogel, A., and Ronacher, B. (2010). Intensity invariance properties of auditory neurons compared to the statistics of relevant natural signals in grasshoppers. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 196(4):285–297.

Cover, T. and Thomas, J. (1991). *Elements of information theory*, volume 1. Wiley Online Library.

Creutzig, F., Benda, J., Wohlgemuth, S., Stumpner, A., Ronacher, B., and Herz, A. V. M. (2010). Timescale-Invariant Pattern Recognition by Feedforward Inhibition and Parallel Signal Processing. *Neural Computation*, 22(6):1493–1510.

Creutzig, F., Wohlgemuth, S., Stumpner, A., Benda, J., Ronacher, B., and Herz, A. V. M. (2009). Timescale-Invariant Representation of Acoustic Communication Signals by a Bursting Neuron. *Journal of Neuroscience*, 29(8):2575–2580.

Desbordes, G., Jin, J., Weng, C., Lesica, N. A., Stanley, G. B., and Alonso, J.-M. (2008). Timing Precision in Population Coding of Natural Scenes in the Early Visual System. *PLoS Biology*, 6(12):e324+.

DeWeese, M. R., Wehr, M., and Zador, A. M. (2003). Binary spiking in auditory cortex. *The Journal of Neuroscience*, 23(21):7940–9.

Dimitrov, A. G., Alexander, Gedeon, T., and Tomas (2006). Effects of stimulus transformations on estimates of sensory neuron selectivity. *Journal of Computational Neuroscience*, 20(3):265–283.

Dörscheidt, G. and Rheinlaender, J. (1980). Computer generation of sound models for behavioural and neurophysiological experiments in insects. *Journal of Insect Physiology*, 26:717–727.

Eggermont, J. J. (2010). Context dependence of spectro-temporal receptive fields with implications for neural coding. *Hearing Research*, 271(1-2):123–132.

Einhäupl, A., Stange, N., Hennig, R. M., and Ronacher, B. (2011). Attractiveness of grasshopper songs correlates with their robustness against noise. *Behavioral Ecology*, 22(4):791–799.

Elsner, N. (1974). Neuroethology of sound production in gomphocerine grasshoppers (Orthoptera: Acrididae). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, V88(1):67–102.

Fairhall, A. L., Burlingame, A. C., Narasimhan, R., Harris, R. A., Puchalla, J. L., and Berry, M. J. (2006). Selectivity for Multiple Stimulus Features in Retinal Ganglion Cells. *Journal of Neurophysiology*, 96(5):2724–2738.

Fairhall, A. L., Lewen, G. D., Bialek, W., and Ruyter, D. (2001). Efficiency and ambiguity in an adaptive neural code. *Nature*, 412(6849):787–792.

Farkhooi, F., Muller, E., and Nawrot, M. P. (2009). Sequential sparsing by successive adapting neural populations. *BMC Neuroscience*, 10(Suppl 1):O10.

Faulkes, Z. and Pollack, G. S. (2000). Effects of Inhibitory Timing on Contrast Enhancement in Auditory Circuits in Crickets (Teleogryllus oceanicus). *Journal of Neurophysiology*, 84(3):1247–1255.

Felsen, G. and Dan, Y. (2005). A natural approach to studying vision. *Nature Neuroscience*, 8(12):1643–1646.

Fischer, B. J. and Westover, M. (2003). The neural multiple access channel. *Neurocomputing*, 52-54:511–518.

Fontaine, B., Goodman, D. F. M., Benichoux, V., and Brette, R. (2011). Brian hears: online auditory processing using vectorization over channels. *Frontiers in Neuroinformatics*, 5:9.

Fox, J. L., Fairhall, A. L., and Daniel, T. L. (2010). Encoding properties of haltere neurons enable motion feature detection in a biological gyroscope. *Proceedings of the National Academy of Sciences*, 107(8):3840–3845.

Fritz, J. B., Shamma, S. A., Elhilali, M., and Klein, D. J. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience*, 6(11):1216–1223.

Gabbiani, F., Krapp, H. G., Koch, C., and Laurent, G. (2002). Multiplicative computation in a visual neuron sensitive to looming. *Nature*, 420(6913):320–4.

Gawne, T. J. and Martin, J. M. (2002). Responses of Primate Visual Cortical V4 Neurons to Simultaneously Presented Stimuli. *Journal of Neurophysiology*, 88:1128–1135.

Geffen, M. N., Broome, B. M., Laurent, G., and Meister, M. (2009). Neural Encoding of Rapidly Fluctuating Odors. *Neuron*, 61(4):570–586.

Geisler, W. S., Najemnik, J., and Ing, A. D. (2009). Optimal stimulus encoders for natural tasks. *Journal of Vision*, 9(13):1–16.

George, a. a., Lyons-Warren, a. M., Ma, X., and Carlson, B. a. (2011). A Diversity of Synaptic Filters Are Created by Temporal Summation of Excitation and Inhibition. *Journal of Neuroscience*, 31(41):14721–14734.

Gollisch, T., Schütze, H., Benda, J., and Herz, A. V. M. (2002). Energy integration describes sound-intensity coding in an insect auditory system. *The Journal of Neuroscience*, 22(23):10434–48.

Gottsberger, B. and Mayer, F. (2007). Behavioral sterility of hybrid males in acoustically communicating grasshoppers (Acrididae, Gomphocerinae). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 193(7):703–714.

Gourévitch, B., Noreña, A., Shaw, G., and Eggermont, J. J. (2009). Spectrotemporal receptive fields in anesthetized cat primary auditory cortex are context dependent. *Cerebral Cortex*, 19(6):1448–61.

Hallem, E. a. and Carlson, J. R. (2006). Coding of odors by a receptor repertoire. *Cell*, 125(1):143–60.

Hardt, M. and Watson, A. H. D. (1994). Distribution of synapses on two ascending interneurones carrying frequency-specific information in the auditory system of the cricket: Evidence for gabaergic inputs. *The Journal of Comparative Neurology*, 345(4):481–495.

Hasenstaub, A. R. and Callaway, E. M. (2010). Paint it black (or red, or green): optical and genetic tools illuminate inhibitory contributions to cortical circuit function. *Neuron*, 67(5):681–4.

Hennig, R. M. (1988). Ascending auditory interneurons in the cricket Teleogryllus commodus (Walker): comparative physiology and direct connections with afferents. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 163(1):135–143.

Hennig, R. M. (2003). Acoustic feature extraction by cross-correlation in crickets? *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 189(8):589–598.

Hennig, R. M. (2009). Walking in Fourier's space: algorithms for the computation of periodicities in song patterns by the cricket Gryllus bimaculatus. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 195(10):971–987.

Hennig, R. M., Franz, A., and Stumpner, A. (2004). Processing of auditory information in insects. *Microscopical Research and Technology*, 63(6):351–374.

Hildebrandt, J. K. (2010). *Neural adaptation in the auditory pathway of crickets and grasshoppers*. Phd thesis, Humboldt-Universität zu Berlin.

Hildebrandt, J. K., Benda, J., and Hennig, R. M. (2009). The origin of adaptation in the auditory pathway of locusts is specific to cell type and function. *The Journal of Neuroscience*, 29(8):2626–36.

Hildebrandt, J. K., Benda, J., and Hennig, R. M. (2011). Multiple Arithmetic Operations in a Single Neuron: The Recruitment of Adaptation Processes in the Cricket Auditory Pathway Depends on Sensory Context. *Journal of Neuroscience*, 31(40):14142–14150.

Hirtz, R. and Wiese, K. (1997). Ultrastructure of synaptic contacts between identified neurons of the auditory pathway in Gryllus bimaculatus, DeGeer. *The Journal of Comparative Neurology*, 386(3):347–357.

Houghton, C. (2009). Studying spike trains using a van Rossum metric with a synapse-like filter. *Journal of Computational Neuroscience*, 26(1):149–155.

Houghton, C. and Sen, K. (2008). A New Multineuron Spike Train Metric. *Neural Computation*, 20(6):1495–1511.

Imaizumi, K. and Pollack, G. S. (1999). Neural coding of sound frequency by cricket auditory receptors. *The Journal of Neuroscience*, 19(4):1508–16.

Imaizumi, K. and Pollack, G. S. (2001). Neural representation of sound amplitude by functionally different auditory receptors in crickets. *The Journal of the Acoustical Society of America*, 109(3):1247–1260.

Imaizumi, K. and Pollack, G. S. (2005). Central projections of auditory receptor neurons of crickets. *The Journal of Comparative Neurology*, 493(3):439–447.

Isaacson, J. S. and Scanziani, M. (2011). How Inhibition Shapes Cortical Activity. *Neuron*, 72(2):231–243.

Ito, I., Ong, R. C.-Y., Raman, B., and Stopfer, M. (2008). Sparse odor representation and olfactory learning. *Nature Neuroscience*, 11(10):1177–84.

Jäkel, F., Schölkopf, B., and Wichmann, F. A. (2009). Does cognitive science need kernels? *Trends in Cognitive Sciences*, 13(9):381–8.

Jia, H., Rochefort, N. L., Chen, X., and Konnerth, A. (2010). Dendritic organization of sensory input to cortical neurons in vivo. *Nature*, 464(7293):1307–1312.

Kalmring, K. (1975). The afferent auditory pathway in the ventral cord of Locusta migratoria (Acrididae). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 104:143–159.

Kapfer, C., Glickfeld, L. L., Atallah, B. V., and Scanziani, M. (2007). Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nature Neuroscience*, 10(6):743–53.

Kaschube, M., Schnabel, M., Löwel, S., Coppola, D. M., White, L. E., and Wolf, F. (2010). Universality in the Evolution of Orientation Columns in the Visual Cortex. *Science*, 330:1113–1116.

Kerlin, A. M., Andermann, M. L., Berezovskii, V. K., and Reid, R. C. (2010). Broadly Tuned Response Properties of Diverse Inhibitory Neuron Subtypes in Mouse Visual Cortex. *Neuron*, 67(5):858–871.

Kim, A. J., Lazar, A. a., and Slutskiy, Y. B. (2011). System identification of Drosophila olfactory sensory neurons. *Journal of Computational Neuroscience*, 30(1):143–61.

Kostarakos, K., Hartbauer, M., and Römer, H. (2008). Matched filters, mate choice and the evolution of sexually selected traits. *PloS one*, 3(8):e3005+.

Kostarakos, K., Hennig, R. M., and Romer, H. (2009). Two matched filters and the evolution of mating signals in four species of cricket. *Frontiers in Zoology*, 6(1):22+.

Krahe, R., Budinger, E., and Ronacher, B. (2002). Coding of a sexually dimorphic song feature by auditory interneurons of grasshoppers: the role of leading inhibition. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 187(12):977–985.

Kumar, A., Rotter, S., and Aertsen, A. (2010). Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nature Reviews Neuroscience*, 11(9):615–627.

Lampl, I., Ferster, D., Poggio, T., and Riesenhuber, M. (2004). Intracellular measurements of spatial integration and the MAX operation in complex cells of the cat primary visual cortex. *Journal of Neurophysiology*, 92(5):2704–13.

Laughlin, S. B. (1981). A simple coding procedure enhances a neuron's information capacity. *Zeitschrift für Naturforschung. Section C: Biosciences*, 36(9-10):910–2.

Laurent, G. (2002). Olfactory network dynamics and the coding of multidimensional signals. *Nature Reviews Neuroscience*, 3(11):884–895.

Leonardo, A. (2005). Degenerate coding in neural systems. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 191(11):995–1010.

Lundstrom, B. N. and Fairhall, A. L. (2006). Decoding Stimulus Variance from a Distributional Neural Code of Interspike Intervals. *Journal of Neuroscience*, 26(35):9030–9037.

Lundstrom, B. N., Higgs, M. H., Spain, W. J., and Fairhall, A. L. (2008). Fractional differentiation by neocortical pyramidal neurons. *Nature Neuroscience*, 11(11):1335–1342.

Luo, S. X., Axel, R., and Abbott, L. F. (2010). Generating sparse and selective third-order responses in the olfactory system of the fly. *Proceedings of the National Academy of Sciences*, 107(23):10713–10718.

Machens, C. K., Schütze, H., Franz, A., Kolesnikova, O., Stemmler, M., Ronacher, B., and Herz, A. V. M. (2003). Single auditory neurons rapidly discriminate conspecific communication signals. *Nature Neuroscience*, 6(4):341–342.

Machens, C. K., Stemmler, M., Prinz, P., Krahe, R., Ronacher, B., and Herz, A. V. M. (2001). Representation of acoustic communication signals by insect auditory receptor neurons. *Journal of Neuroscience*, 21(9):3215–3227.

Machens, C. K., Wehr, M., and Zador, A. M. (2004). Linearity of cortical receptive fields measured with natural sounds. *Journal of Neuroscience*, 24(5):1089–1100.

Margoliash, D. and Konishi, M. (1985). Auditory representation of autogenous song in the song system of white-crowned sparrows. *Proceedings of the National Academy of Sciences of the United States of America*, 82(17):5997–6000.

Marsat, G. and Pollack, G. S. (2004). Differential temporal coding of rhythmically diverse acoustic signals by a single interneuron. *Journal of Neurophysiology*, 92(2):939–948.

Marsat, G. and Pollack, G. S. (2005). Effect of the Temporal Pattern of Contralateral Inhibition on Sound Localization Cues. *Journal of Neuroscience*, 25(26):6137–6144.

Marsat, G. and Pollack, G. S. (2006). A Behavioral Role for Feature Detection by Sensory Bursts. *Journal of Neuroscience*, 26(41):10542–10547.

Marsat, G. and Pollack, G. S. (2007). Efficient inhibition of bursts by bursts in the auditory system of crickets. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 193(6):625–633.

Marsat, G. and Pollack, G. S. (2010). The structure and size of sensory bursts encode stimulus information but only size affects behavior. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 196(4):315–320.

McCotter, M., Gosselin, F., Sowden, P., and Schyns, P. (2005). The use of visual information in natural scenes. *Visual Cognition*, 12(6):938–953.

Meyer, J. and Elsner, N. (1996). How well are frequency sensitivities of grasshopper ears tuned to species-specific song spectra? *Journal of Experimental Biology*, 199:1631–1642.

Meyer, J. and Elsner, N. (1997). Can spectral cues contribute to species separation in closely related grasshoppers? *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 180(2):171–180.

Mineault, P. J., Barthelmé, S., and Pack, C. C. (2009). Improved classification images with sparse priors in a smooth basis. *Journal of Vision*, 9(10):1–24.

Mitchell, M. (1998). *An introduction to genetic algorithms*. The MIT press.

Murray, R. F. (2011). Classification images: A review. *Journal of Vision*, 11(5):1–25.

Narayan, R., Best, V., Ozmeral, E., Mcclaine, E., Dent, M., Cunningham, B. S., and Sen, K. (2007). Cortical interference effects in the cocktail party problem. *Nature Neuroscience*, 10(12):1601–1607.

Neri, P. (2004). Estimation of nonlinear psychophysical kernels. *Journal of Vision*, 4(2):82–91.

Neri, P. and Levi, D. M. (2006). Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vision Research*, 46(16):2465–2474.

Neuhofer, D., Stemmler, M., and Ronacher, B. (2011). Neuronal precision and the limits for acoustic signal recognition in a small neuronal network. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral physiology*, 197(3):251–65.

Neuhofer, D., Wohlgemuth, S., Stumpner, A., and Ronacher, B. (2008). Evolutionarily conserved coding properties of auditory neurons across grasshopper species. *Proceedings of the Royal Society of London. Series B*, 275(1646):1965–1974.

Nolen, T. G. and Hoy, R. (1984). Initiation of behavior by single neurons: the role of behavioral context. *Science*, 226(4677):992–994.

Nolen, T. G. and Hoy, R. (1986a). Phonotaxis in flying crickets. I. Attraction to the calling song and avoidance of bat-like ultrasound are discrete behaviors. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 159(4):423–439.

Nolen, T. G. and Hoy, R. (1986b). Phonotaxis in flying crickets. II. Physiological mechanisms of two-tone suppression of the high frequency avoidance steering behavior by the calling song. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 159(4):441–456.

Nolen, T. G. and Hoy, R. (1987). Postsynaptic inhibition mediates high-frequency selectivity in the cricket Teleogryllus oceanicus: implications for flight phonotaxis behavior. *The Journal of Neuroscience*, 7(7):2081–2096.

Ohiorhenuan, I. E., Mechler, F., Purpura, K. P., Schmid, A. M., Hu, Q., and Victor, J. D. (2010). Sparse coding and high-order correlations in fine-scale cortical networks. *Nature*, 466(7306):617–621.

Olsen, S. R., Bhandawat, V., and Wilson, R. I. (2010). Divisive normalization in olfactory population codes. *Neuron*, 66(2):287–99.

Olsen, S. R. and Wilson, R. I. (2008). Lateral presynaptic inhibition mediates gain control in an olfactory circuit. *Nature*, 452(7190):956–60.

Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609.

Olshausen, B. A. and Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487.

Osborne, L. C., Palmer, S. E., Lisberger, S. G., and Bialek, W. (2008). The Neural Basis for Combinatorial Coding in a Cortical Population Response. *Journal of Neuroscience*, 28(50):13522–13531.

Papadopoulou, M., Cassenaer, S., Nowotny, T., and Laurent, G. (2011). Normalization for Sparse Encoding of Odors by a Wide-Field Interneuron. *Science*, 332(6030):721–725.

Perez-Orive, J., Mazor, O., Turner, G. C., Cassenaer, S., Wilson, R. I., and Laurent, G. (2002). Oscillations and sparsening of odor representations in the mushroom body. *Science*, 297(5580):359–65.

Petersen, R. S., Brambilla, M., Bale, M. R., Alenda, A., Panzeri, S., Montemurro, M. A., and Maravall, M. (2008). Diverse and Temporally Precise Kinetic Feature Selectivity in the VPm Thalamic Nucleus. *Neuron*, 60(5):890–903.

Pflüger, H. and Field, L. (1999). A locust chordotonal organ coding for proprioceptive and acoustic stimuli. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 184:169–183.

Pienkowski, M. and Eggermont, J. J. (2011). Sound frequency representation in primary auditory cortex is level tolerant for moderately loud, complex sounds. *Journal of Neurophysiology*, 106(2):1016–27.

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., and Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995–999.

Pillow, J. W. and Simoncelli, E. P. (2006). Dimensionality reduction in neural models: An information-theoretic generalization of spike-triggered average and covariance analysis. *Journal of Vision*, 6(4):414–428.

Pollack, G. S. (1988). Selective attention in an insect auditory neuron. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 8(7):2635–2639.

Pollack, G. S. (1994). Synaptic inputs to the omega neuron of the cricket Teleogryllus oceanicus: differences in EPSP waveforms evoke by low and high sound frequencies. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 174(1):83–89.

Poo, C. and Isaacson, J. S. (2009). Odor representations in olfactory cortex: "sparse" coding, global inhibition, and oscillations. *Neuron*, 62(6):850–861.

Prescott, S. a. and De Koninck, Y. (2003). Gain control of firing rate by shunting inhibition: roles of synaptic noise and dendritic saturation. *Proceedings of the National Academy of Sciences*, 100(4):2076–81.

Priebe, N. J. and Ferster, D. (2008). Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron*, 57(4):482–97.

Quiroga, R. Q. and Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nature Reviews Neuroscience*, 10(3):173–185.

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature*, 435(7045):1102–7.

Riffell, J. A., Lei, H., Christensen, T. A., and Hildebrand, J. G. (2009). Characterization and coding of behaviorally significant odor mixtures. *Current Biology*, 19(4):335–40.

Robert, D. (1989). The Auditory Behaviour of Flying Locusts. *Journal of Experimental Biology*, 147(1):279–301.

Roddey, C. J., Girish, B., and Miller, J. P. (2000). Assessing the Performance of Neural Encoding Models in the Presence of Noise. *Journal of Computational Neuroscience*, V8(2):95–112.

Rokem, A., Watzl, S., Gollisch, T., Stemmler, M., Herz, A. V. M., and Samengo, I. (2006). Spike-timing precision underlies the coding efficiency of auditory receptor neurons. *Journal of Neurophysiology*, 95(4):2541–2552.

Römer, H. and Marquart, V. (1984). Morphology and physiology of auditory interneurons in the metathoracic ganglion of the locust. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, V155(2):249–262.

Römer, H. and Seikowski, U. (1985). Responses to model songs of auditory neurons in the thoracic ganglia and brain of the locust. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, V156(6):845–860.

Ronacher, B. and Hennig, R. M. (2004). Neuronal adaptation improves the recognition of temporal patterns in a grasshopper. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 190(4):311–319.

Ronacher, B. and Stumpner, A. (1988). Filtering of behaviourally relevant temporal parameters of a grasshopper's song by an auditory interneuron. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 163(4):517–523.

Rothman, J. S., Cathala, L., Steuber, V., and Silver, A. R. (2009). Synaptic depression enables neuronal gain control. *Nature*, 457(7232):1015–1018.

Rozell, C. J., Johnson, D. H., Baraniuk, R. G., and Olshausen, B. a. (2008). Sparse coding via thresholding and local competition in neural circuits. *Neural Computation*, 20(10):2526–63.

Runyan, C. A., Schummers, J., Wart, A. V., Kuhlman, S. J., Wilson, N. R., Huang, Z. J., Sur, M., and Van Wart, A. (2010). Response Features of Parvalbumin-Expressing Interneurons Suggest Precise Roles for Subtypes of Inhibition in Visual Cortex. *Neuron*, 67(5):847–857.

Rust, N. C., Schwartz, O., Movshon, J. A., and Simoncelli, E. P. (2005). Spatiotemporal elements of macaque v1 receptive fields. *Neuron*, 46(6):945–56.

Schildberger, K. (1984). Temporal selectivity of identified auditory neurons in the cricket brain. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 155(2):171–185.

Schildberger, K. (1988). Behavioral and neuronal mechanisms of cricket phonotaxis. *Cellular and Molecular Life Sciences*, 44(5):408–415.

Schildberger, K., Milde, J. J., and Hörner, M. (1988). The function of auditory neurons in cricket phonotaxis I. Influence of hyperpolarization of identified neurons on sound localization. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 163(5):633–640.

Schmidt, A., Riede, K., and Römer, H. (2011). High background noise shapes selective auditory filters in a tropical cricket. *The Journal of Experimental Biology*, 214(Pt 10):1754–62.

Schmidt, A., Ronacher, B., and Hennig, R. M. (2008). The role of frequency, phase and time for processing of amplitude modulated signals by grasshoppers. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 194(3):221–233.

Schmuker, M. and Schneider, G. (2007). Processing and classification of chemical data inspired by insect olfaction. *Proceedings of the National Academy of Sciences*, 104(51):20285–9.

Schneider, D. M. and Woolley, S. M. (2010). Discrimination of Communication Vocalizations by Single Neurons and Groups of Neurons in the Auditory Midbrain. *Journal of Neurophysiology*, 103(6):3248–3265.

Schneider, D. M. and Woolley, S. M. (2011). Extra-Classical Tuning Predicts Stimulus-Dependent Receptive Fields in Auditory Neurons. *Journal of Neuroscience*, 31(33):11867–11878.

Schneidman, E., Puchalla, J. L., Segev, R., Harris, R. a., Bialek, W., and Berry, M. J. (2011). Synergy from Silence in a Combinatorial Neural Code. *Journal of Neuroscience*, 31(44):15732–15741.

Schreiber, S., Fellous, J.-M., Whitmer, D., Tiesinga, P. H., and Sejnowski, T. J. (2003). A New Correlation-Based Measure of Spike Timing Reliability. *Neurocomputing*, 52-54:925–931.

Schul, J., Holderied, M., von Helversen, D., and von Helversen, O. (1999). Directional hearing in grasshoppers: neurophysiological testing of a bioacoustic model. *Journal of Experimental Biology*, 202 (Pt 2):121–133.

Schwartz, O., Pillow, J. W., Rust, N. C., and Simoncelli, E. P. (2006). Spike-triggered neural characterization. *Journal of Vision*, 6(4):484–507.

Selverston, A. I., Kleindienst, H., and Huber, F. (1985). Synaptic connectivity between cricket auditory interneurons as studied by selective photoinactivation. *Journal of Neuroscience*, 5(5):1283–1292.

Seriès, P., Latham, P. E., and Pouget, A. (2004). Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nature Neuroscience*, 7(10):1129–1135.

Seriès, P., Stocker, A., and Simoncelli, E. P. (2009). Is the Homunculus "Aware" of Sensory Adaptation? *Neural Computation*, 21(12):3271–3304.

Sharpee, T. O., Nagel, K. I., and Doupe, A. J. (2011). Two-dimensional adaptation in the auditory forebrain. *Journal of Neurophysiology*, 106(4):1841–61.

Sharpee, T. O., Rust, N. C., and Bialek, W. (2004). Analyzing neural responses to natural signals: maximally informative dimensions. *Neural Computation*, 16(2):223–250.

Sharpee, T. O., Sugihara, H., Kurgansky, A. V., Rebrik, S. P., Stryker, M. P., and Miller, K. D. (2006). Adaptive filtering enhances information transmission in visual cortex. *Nature*, 439(7079):936–42.

Shub, D. E. and Richards, V. M. (2009). Psychophysical spectro-temporal receptive fields in an auditory task. *Hearing Research*, 251(1-2):1–9.

Silbering, A. F. and Galizia, C. G. (2007). Processing of odor mixtures in the Drosophila antennal lobe reveals both global inhibition and glomerulus-specific interactions. *The Journal of Neuroscience*, 27(44):11966–77.

Skiebe, P., Corrette, B. J., and Wiese, K. (1990). Evidence that histamine is the inhibitory transmitter of the auditory interneuron ON1 of crickets. *Neuroscience letters*, 116(3):361–6.

Slee, S. J., Higgs, M. H., Fairhall, A. L., and Spain, W. J. (2005). Two-dimensional time coding in the auditory brainstem. *Journal of Neuroscience*, 25(43):9978–9988.

Smith, E. C. and Lewicki, M. S. (2006). Efficient auditory coding. *Nature*, 439(7079):978–982.

Sokoliuk, T., Stumpner, A., and Ronacher, B. (1989). GABA-like immunoreactivity suggests an inhibitory function of the thoracic low-frequency neuron (TN1) in acridid grasshoppers. *Naturwissenschaften*, 76(5):223–225.

Stopfer, M., Bhagavan, S., Smith, B. H., and Laurent, G. (1997). Impaired odour discrimination on desynchronization of odour-encoding neural assemblies. *Nature*, 390(6655):70–4.

Stumpner, A. (1988). *Auditorische thorakale Interneurone von Chorthippus biguttulis L.: Morphologische und physiologische Charakterisierung und Darstellung ihrer Filtereigenschaften für verhaltensrelevante Lautattrappen*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg.

Stumpner, A. (1989). Physiological variability of auditory neurons in a grasshopper. *Naturwissenschaften*, 76(9):427–429.

Stumpner, A. and Helversen, O. (1992). Recognition of a two-element song in the grasshopper Chorthippus dorsatus (Orthoptera: Gomphocerinae). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 171(3):405–412.

Stumpner, A. and Ronacher, B. (1991). Auditory Interneurones in the Metathoracic Ganglion of the Grasshopper Chorthippus Biguttulus: I. Morphological and Physiological Characterization. *Journal of Experimental Biology*, 158(1):391–410.

Stumpner, A. and Ronacher, B. (1994). Neurophysiological Aspects of Song Pattern Recognition and Sound Localization in Grasshoppers. *American Zoologist*, 34(6):696–705.

Stumpner, A., Ronacher, B., and von Helversen, O. (1991). Auditory Interneurones in the Metathoracic Ganglion of the Grasshopper Chorthippus Biguttulus: II. Processing of Temporal Patterns of the Song of the Male. *Journal of Experimental Biology*, 158(1):411–430.

Theunissen, F. E. and Miller, J. P. (1995). Temporal encoding in nervous systems: a rigorous definition. *Journal of Computational Neuroscience*, 2(2):149–162.

Tripp, B. P. and Eliasmith, C. (2010). Population Models of Temporal Differentiation. *Neural Computation*, 22(3):621–659.

Tunstall, D. N. and Pollack, G. S. (2005). Temporal and directional processing by an identified interneuron, ON1, compared in cricket species that sing with different tempos. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 191(4):363–372.

Uchida, N. and Mainen, Z. F. (2008). Odor concentration invariance by chemical ratio coding. *Frontiers in Neuroscience - Computational Neuroscience*.

van Rossum, M. C. W. (2001). A novel spike distance. *Neural Computation*, 13(4):751–763.

van Wezel, R. J., Lankheet, M. J., Verstraten, F. A., Maree, A. F., and van de Grind, W. A. (1996). Responses of complex cells in area 17 of the cat to bi-vectorial transparent motion. *Vision research*, 36(18):2805–13.

Vedenina, V. Y., Panyutin, A. K., and Von (2007). The unusual inheritance pattern of the courtship songs in closely related grasshopper species of the Chorthippus albomarginatus-group (Orthoptera: Gomphocerinae). *Journal of Evolutionary Biology*, 20(1):260–277.

Victor, J. D. and Purpura, K. P. (1997). Metric-space analysis of spike trains: theory, algorithms and application. *Network: Computation in Neural Systems*, 8(2):127–164.

Viemeister, N. F. and Wakefield, G. H. (1991). Temporal integration and multiple looks. *The Journal of the Acoustical Society of America*, 90(2):858–865.

Vinje, W. E. and Gallant, J. L. (2000). Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science*, 287(5456):1273–1276.

Vogel, A., Hennig, R. M., and Ronacher, B. (2005). Increase of neuronal response variability at higher processing levels as revealed by simultaneous recordings. *Journal of Neurophysiology*, 93(6):3548–3559.

Vogel, A. and Ronacher, B. (2007). Neural Correlations Increase Between Consecutive Processing Levels in the Auditory System of Locusts. *Journal of Neurophysiology*, 97(5):3376–3385.

von Helversen, D. (1972). Gesang des M\"{a}nnchens und Lautschema des Weibchens bei der Feldheuschrecke Chorthippus biguttulus (Orthoptera, Acrididae). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 81(4):381–422.

von Helversen, D. (1997). Acoustic communication and orientation in grasshoppers. In Lehrer, M., editor, *Orientation and communication in arthropods*, pages 301–342. Birkhäuser Basel.

von Helversen, D. and von Helversen, O. (1997). Recognition of sex in the acoustic communication of the grasshopper Chorthippus biguttulus (Orthoptera, Acrididae). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 180(4):373–386.

von Helversen, D. and von Helversen, O. (1998). Acoustic pattern recognition in a grasshopper: processing in the time or frequency domain? *Biological Cybernetics*, 79(6):467–476.

von Helversen, O. and von Helversen, D. (1987). Innate receiver mechanisms in the acoustic communication of orthopteran insects. In Guthrie, D. M., editor, *Aims and methods in neuroethology*. Manchester University Press.

Wang, H.-P., Spencer, D., Fellous, J.-M., and Sejnowski, T. J. (2010). Synchrony of Thalamocortical Inputs Maximizes Cortical Reliability. *Science*, 328(5974):106–109.

Wang, X.-J., Liu, Y., Vives, S. M. V., and McCormick, D. A. (2003). Adaptation and Temporal Decorrelation by Single Neurons in the Primary Visual Cortex. *Journal of Neurophysiology*, 89(6):3279–3293.

Wehr, M. and Zador, A. M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature*, 426(6965):442–6.

Weschke, G. and Ronacher, B. (2008). Influence of sound pressure level on the processing of amplitude modulations by auditory neurons of the locust. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 194(3):255–265.

Wiechert, M. T., Judkewitz, B., Riecke, H., and Friedrich, R. W. (2010). Mechanisms of pattern decorrelation by recurrent neuronal circuits. *Nature Neuroscience*, 13(8):1003–1010.

Wilent, W. B. and Contreras, D. (2005). Dynamics of excitation and inhibition underlying stimulus selectivity in rat somatosensory cortex. *Nature Neuroscience*, 8(10):1364–70.

Willmore, B. D., Mazer, J., and Gallant, J. L. (2011). Sparse coding in striate and extrastriate visual cortex. *Journal of Neurophysiology*, 105(6):2907–19.

Willmore, B. D. and Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Network: Computation in Neural Systems*, 12(3):255–70.

Wimmer, K., Hildebrandt, J. K., Hennig, R. M., and Obermayer, K. (2008). Adaptation and Selective Information Transmission in the Cricket Auditory Neuron AN2. *PLoS Computational Biology*, 4(9):e1000182+.

Wittmann, J. P., Kolss, M., and Reinhold, K. (2011). A neural network-based analysis of acoustic courtship signals and female responses in Chorthippus biguttulus grasshoppers. *Journal of Computational Neuroscience*, 31(1):105–15.

Wohlers, D. and Huber, F. (1982). Processing of sound signals by six types of neurons in the prothoracic ganglion of the cricket, Gryllus campestris L. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 146(2):161–173.

Wohlgemuth, S. and Ronacher, B. (2007). Auditory Discrimination of Amplitude Modulations Based on Metric Distances of Spike Trains. *Journal of Neurophysiology*, 97(4):3082–3092.

Wu, G. K., Arbuckle, R., Liu, B.-H., Tao, H. W., and Zhang, L. I. (2008). Lateral Sharpening of Cortical Frequency Tuning by Approximately Balanced Inhibition. *Neuron*, 58(1):132–143.

Wyttenbach, R. A., May, M. L., and Hoy, R. (1996). Categorical Perception of Sound Frequency by Crickets. *Science*, 273(5281):1542–1544.

Xing, D., Ringach, D. L., Hawken, M., and Shapley, R. M. (2011). Untuned Suppression Makes a Major Contribution to the Enhancement of Orientation Selectivity in Macaque V1. *Journal of Neuroscience*, 31(44):15972–15982.

Yaksi, E., Saint, V., Niessing, J. J., Bundschuh, S. T., Friedrich, R. W., and von Saint Paul, F. (2009). Transformation of odor representations in target areas of the olfactory bulb. *Nature Neuroscience*, 12(4):474–482.

Zhaoping, L. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, 6(1):9–16.

Zhaoping, L. (2006). Theoretical understanding of the early visual processes by data compression and data selection. *Network: Computation in Neural Systems*, 17(1):301–334.

Zilany, M. S., Bruce, I. C., Nelson, P. C., and Carney, L. H. (2009). A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics. *The Journal of the Acoustical Society of America*, 126(5):2390–412.

Zoccolan, D., Cox, D., and DiCarlo, J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *The Journal of Neuroscience*, 25(36):8150–64.

Danksagung

Ich möchte Bernhard Ronacher dafür danken, dass er mir durch sein Vertrauen so viel Freiraum bei der Gestaltung dieser Arbeit gegeben hat. Er war immer da wenn ich Probleme oder Fragen hatte, konnte mir mit seinem umfangreichen Wissen über unsere Modellorganismen immer weiterhelfen. Desweiteren hat er geholfen meine Texte verständlicher und fehlerfreier zu machen.

Matthias Hennig möchte ich für seine Enthusiasmus danken - ohne seine Begeisterung wäre das Kapitel über Grillen sicher nie entstanden. Auch er hat sich immer Zeit genommen, wenn ich Fragen hatte. Ohne die von ihm entwickelten Programme wäre die Aufnahme von elektrophysiologischen und Verhaltensdaten weniger reibungslos verlaufen.

Susanne Schreiber möchte ich für die sehr schöne Zusammenarbeit und die Ermutigung danken - sie hat mir vor allem zum Ende der Arbeit sehr viel Ansporn gegeben.

Sandra Wohlgemuth hat mich zur Beschäftigung mit der Populationskodierung bei Grasshüpfern ermutigt und mir einen umfangreichen Datensatz dazu überlassen, der viele Hypothesen, die in Teil 1 der Arbeit eingeflossen sind, inspiriert hat. Ich danke ihr dafür.

Ich danke all jenen, die Daten zu dieser Arbeit beigetragen haben - ohne sie wäre ein Großteil dieser Arbeit nicht möglich gewesen. Neben Sandra Wohlgemuth haben Olaf Kutzki, Florian Rau und Viktor Naumov elektrophysiologische Daten gesammelt. Olaf Kutzki, Nicole Stange, Jana Sträter, Anneke Fliß und Stefanie Krämer haben mir ihre Verhaltensdaten überlassen. Ich weiss wieviel Mühe das Sammeln von Daten macht und bin euch sehr dankbar dafür!

Neben Susanne Schreiber danke ich auch Felix Wichmann und Martin Nawrot, die als Mitglieder meines Kommitees meine Arbeit mit wichtigen Kommentaren und Anregungen begleitet haben. Die gemeinsamen Retreats mit Martin Nawrot und die Diskussionen mit ihm und den Mitgliedern seiner AG haben mein Denken über neuronale Kodierung geprägt.

Viele Gespräche mit Conor Houghton haben mein Denken über "sparse coding" geprägt. Auch Jan Benda's Interesse an meiner Arbeit und seine Bereitschaft auf Konferenzen Fragen zu meinem Poster zu beantworten haben mir sehr geholfen.

Ich möchte Jan-Hendrik Schleimer danken, der in vielen Diskussion mit seinem umfangreichen theoretischen Wissen unglaublich geholfen hat. Ohne die Diskussionen mit Vinzenz Schönfelder über Klassifikationsmaschinen hätte ich sicherlich nie den Ansatz, der zu Kapitel 4 geworden ist, probiert. Danke!

Jannis Hildebrandt war zu Beginn der Dissertation mein Zimmergenosse - ohne ihn hätte ich die dB-Skala wahrscheinlich nie verstanden. Florian Rau hat das Zimmer mit mir zum Ende der Doktorarbeit geteilt - ohne ihn hätte die Zeit nur halb so viel Spaß gemacht.

Ich danke Olaf dafür, dass er immer ein Würstchen für mich übrig hatte. Und natürlich

auch für die vielen Diskussion über Grasshüpfer. Sein gesunde Skeptik wann immer ich eine unausgegorene Idee vorgetragen habe, hat mich ungemein angespornt.

Ich danke Frederic Römschied für die vielen Diskussionen über Temperatur und die gemeinsamen Konferenzen.

Ich danke Regina Lübke dafür, dass sie immer für frische Versuchstiere gesorgt hat und immer ein offenes Ohr offen für kleine und große Probleme hatte. Vanessa Cassagrande, Margret Franke und Karin Winkelhöfer haben für ein reibungsloses Funktionieren im Hintergrund gesorgt - ich danke ihnen dafür.

Den Mitgliedern der Abteilung Verhaltensphysiologie danke ich für die vielen gemeinsam verlebten Kaffeepausen und Kicker- und Grillabende.

Arnulf Köhncke danke ich für seine Freundschaft. Mit ihm konnte ich nicht nur über Wissenschaft im Allgemeinen und Speziellen, sondern über alles andere auch reden.

Ich danke meine Eltern für ihre Unterstützung, Geduld, Aufmunterung und ihr Vertrauen.

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 27.03.2012 Clemens, Jan